CANADIAN JOURNAL OF MATHEMATICS

Journal Canadien de Mathématiques

VOL. IX - NO. 4
1957

Published for
THE CANADIAN MATHEMATICAL CONGRESS
by the
University of Toronto Press

The chief languages of the *Journal* are English and French.

Manuscripts for publication in the *Journal* should be sent to the *Editor-in-Chief*, H. S. M. Coxeter, University of Toronto. Everything possible should be done to lighten the task of the reader; the notation and reference system should be carefully thought out. Every paper should contain an introduction summarizing the results as far as possible in such a way as to be understood by the non-expert.

All other correspondence should be addressed to the *Managing Editor*, G. de B. Robinson, University of Toronto.

The *Journal* is published quarterly. Subscriptions should be sent to the *Managing Editor*. The price per volume of four numbers is $8.00. This is reduced to $4.00 for individual members of recognized Mathematical Societies.

The Canadian Mathematical Congress gratefully acknowledges the assistance of the following towards the cost of publishing this Journal:

University of Alberta	Assumption University
University of British Columbia	Carleton College
Dalhousie University	Ecole Polytechnique
Université Laval	Loyola College
University of Manitoba	McGill University
McMaster University	Université de Montréal
Queen's University	Royal Military College
St. Mary's University	University of Toronto

National Research Council of Canada
and the
American Mathematical Society

AUTHORIZED AS SECOND CLASS MAIL, POST OFFICE DEPARTMENT, OTTAWA

THE PROJECTION OF A LINEAR FUNCTIONAL ON THE MANIFOLD OF INTEGRALS

H. GORDON AND E. R. LORCH

I. ON INTEGRALS

Preliminaries. If μ is a measure defined on a space $\mathscr{E}$ and $f_n(x)$ is a sequence of μ-integrable functions converging to $f(x)$, then under suitable conditions of boundedness, one has the theorem of Lebesgue that

$$\lim_{n\to\infty} \int f_n(x)\, d\,\mu(x) = \int f(x)\, d\,\mu(x).$$

In particular,

(1) $$f_n(x) \uparrow f(x) \text{ implies } \int f_n(x) d\mu(x) \to \int f(x)\, d\,\mu(x).$$

Note that the convergence of $\{f_n(x)\}$ to $f(x)$ is pointwise and not uniform. The measure $\mu(x)$ gives rise to a linear functional Ff on the space of μ-integrable functions: $Ff = \int f(x)\, d\,\mu(x)$.

The converse procedure has been studied at length by Daniell. Starting with a suitable collection of functions $f(x)$ where x belongs to an abstract set $\mathscr{E}$, Daniell considers a positive linear functional Ff (that is, $f(x) \geqslant 0$ implies $Ff \geqslant 0$) endowed with the property: $f_n(x) \uparrow f(x)$ implies $Ff_n \to Ff$. He shows that this linear functional has essentially all the properties of an integral —for example, with its help it is possible to extend in the classic fashion the given class of functions $f(x)$ to the class of summable functions. Thus the theorem of Lebesgue and the work of Daniell establish the fact that the existence of an integral is essentially equivalent to the existence of a linear functional with the property:

(2) $$f_n(x) \uparrow f(x) \quad \text{implies} \quad Ff_n \to Ff.$$

In the future when we refer to a positive integral, we shall mean a positive linear functional which has the property (2). When we refer to an arbitrary integral we shall mean a linear functional F which is the difference of two positive integrals: $F = F_1 - F_2$, $F_1 \geqslant 0$, $F_2 \geqslant 0$.

The aim of this note is to establish a decomposition theorem for linear functionals on particular Banach spaces of functions. Thus given an abstract set $\mathscr{E}$ and a Banach space $\mathfrak{B}$ of functions $f(x)$, $x \in \mathscr{E}$, which will be described more fully below, let F represent a bounded linear functional on $\mathfrak{B}$. We shall see that there is a unique decomposition $F = G + H$ where G is an integral and H is purely finitely additive in a sense to be defined. Furthermore, if F

Received December 1, 1955. This work was carried out with the help of a grant from the National Science Foundation. The result contained in this paper was reported in (1).

is positive, $F \geqslant 0$, then also $G \geqslant 0$ and $H \geqslant 0$. If $F \geqslant 0$ and $H \geqslant G$, then $G = 0$.

This theorem may be considered to be an abstract formulation of a result in the theory of measure which has been noted by several authors[1] and which has received its most complete formulation from Yosida and Hewitt **(3)**. The latter consider an additive positive measure ϕ defined on a set X and prove that there exist positive measures ϕ_c, ϕ_p such that ϕ_c is completely additive and ϕ_p is purely finitely additive and such that $\phi = \phi_c + \phi_p$.

The well-known Riesz representation theorem states that every linear functional on the space C of real continuous functions $f(x)$, $0 \leqslant x \leqslant 1$, may be represented as an integral of the form

$$\int_0^1 f(x)\, d\alpha(x).$$

This theorem has been generalized to the following form. Let $\mathscr{R}$ be a compact Hausdorff space and let $C(\mathscr{R})$ be the Banach space of continuous functions $f(x)$ with the norm $\| f \| = \text{l.u.b.} \mid f(x) \mid, x \in \mathscr{R}$. Then if F is a linear functional on $C(\mathscr{R})$, there is a completely additive measure $\mu(x)$ defined over the Borel sets of $\mathscr{R}$ such that

$$Ff = \int_{\mathscr{R}} f(x)\, d\mu(x).$$

In other words, every linear functional is an integral in our sense. However, the latter result is immediate. For if $\{f_n(x)\}$ is a monotone sequence of functions converging to $f(x)$, then since $\mathscr{R}$ is compact, pointwise convergence implies uniform convergence. Thus $f_n(x) \uparrow f(x)$ implies $\| f - f_n \| \to 0$ which in turn implies $Ff_n \to Ff$ for an arbitrary bounded linear F.

Example. We give an example of a linear functional F which does not have property (2) and hence is not an integral.

Let $f(x)$ be a continuous function defined on the real line which vanishes outside a compact set. Define the functional F by $Ff = 0$. The totality of such functions constitutes a linear manifold. Consider the closure $\mathfrak{M}$ of this manifold in the uniform topology; on $\mathfrak{M}$ let F be defined by continuity (to be zero). We adjoin to $\mathfrak{M}$ the set of constants λ and define $F\lambda = \lambda$. Then the system so obtained is a Banach space and F is a bounded linear functional on this space. There exists a sequence $\{f_n(x)\}$ of functions each of which vanish outside a compact set and such that $f_n(x) \uparrow 1$ for all x. Evidently $0 = Ff_n(x) \neq F1 = 1$.

II. The Fundamental Projection

The Space. Let $\mathscr{E}$ be an abstract set. Consider a set $\mathfrak{B}$ of bounded real functions on $\mathscr{E}$. Suppose that $\mathfrak{B}$ has the properties:

[1] Woodbury **(2)** mentions that the result for measures was known to B. Jessen. See also the paper of Yosida and Hewitt **(3**, footnote **2)**, whose decomposition theorem was also known to Kakutani. (*Added in Proof:* See also recent work of H. Bauer, in particular, Math. Z., *65* (1956), 448–482.)

(1) $\mathfrak{B}$ is a vector space — that is, $f(x), g(x) \in \mathfrak{B}$ imply $\alpha f(x) + \beta g(x) \in \mathfrak{B}$ for real α, β;

(2) $\mathfrak{B}$ is a lattice — that is, $f(x)$ and $g(x) \in \mathfrak{B}$ imply that $f(x) \wedge g(x) = \inf (f(x), g(x))$ and $f(x) \vee g(x) = \sup (f(x), g(x)) \in \mathfrak{B}$;

(3) $\mathfrak{B}$ is closed in the uniform topology—that is, $\mathfrak{B}$ contains the uniform limit of any sequence in $\mathfrak{B}$. If we write

$$\| f \| = \text{l.u.b.} \, | f(x) | \, , x \in \mathcal{E}$$

then $\mathfrak{B}$ is a real Banach space. It may be noted that if $\mathfrak{B}$ is an algebra, condition (2) is dependent on the remaining hypotheses.

We shall be interested in the space $\mathfrak{B}^*$ consisting of all bounded linear functionals defined on $\mathfrak{B}$. If the functional $F \in \mathfrak{B}^*$ has the property $Ff \geqslant 0$ for each $f(x) \geqslant 0$ then F is called positive.

The following theorem is proved by methods well known in the theory of measure.

If F is an arbitrary bounded linear functional over $\mathfrak{B}$, there exist two positive functionals G and H such that $F = G - H$. Furthermore G and H are bounded and satisfy $\| G \| \leqslant \| F \|$, $\| H \| \leqslant \| F \|$.

Proof. If $f(x) \in \mathfrak{B}$, then there exist positive functions $g(x), h(x) \in \mathfrak{B}$ such that $f(x) = g(x) - h(x)$. Indeed, we may take $g(x) = f(x) \vee 0$ and $h(x) = (-f(x)) \vee 0$. Since $Ff = F(g - h) = Fg - Fh$, F is completely known if it is known on the cone of positive elements.

Let $f(x) \geqslant 0$ be fixed and define

$$Gf = \underset{0 \leqslant h \leqslant f}{\text{l.u.b.}} \; Fh.$$

Then for positive f, f_1, and f_2 it may be seen that $G(f_1 + f_2) = Gf_1 + Gf_2$ and if $\alpha > 0$, $G\alpha f = \alpha Gf$. It is necessary to use the lattice properties of $\mathfrak{B}$ in order to establish this fact. Thus, if for positive functions f_i, g_i we have $f_1 + g_1 = f_2 + g_2$, then $Gf_1 - Gf_2 = Gg_2 - Gg_1$. The definition of G is completed as follows: If f is arbitrary, let $f = g - h$ where g and h are positive, and set $Gf = Gg - Gh$. It may be seen that G is a linear functional. If $f \geqslant 0$, since

$$Gf = \underset{0 \leqslant h \leqslant f}{\text{l.u.b.}} \; Fh,$$

$Gf \geqslant 0$; that is, G is positive. Finally G is bounded and for the bounds of G and F we have $\|G\| \leqslant \|F\|$. This follows from the inequalities

$$Gf = G(f^+ - f^-) = Gf^+ - Gf^- \leqslant Gf^+ \leqslant \|F\| \, \|f^+\| \leqslant \|F\| \, \|f\|.$$

Similarly $Gf \geqslant - \|F\| \, \|f\|$. Here f^+ and f^- denote the positive and negative parts of f. Thus $\|G\| \leqslant \|F\|$.

The linear functional H is defined by $H = G - F$ or by

$$Hf = \underset{0 \leqslant g \leqslant f}{\text{l.u.b.}} \; (-Fg).$$

It is clear that $\|H\| \leqslant \|F\|$.

We may now prove that the totality of integrals over $\mathfrak{B}$ is a closed linear manifold. If F_1 and F_2 are integrals (that is, each is the difference of two positive integrals), then obviously $c_1F_1 + c_2F_2$ is also an integral. Thus the integrals form a linear manifold. We shall see that this manifold is closed.

Let $\{F_n\}$ be a sequence of integrals and let F be a linear functional such that $\|F_n - F\| \to 0$. Suppose $F_n = G_n - H_n$ where G_n and H_n are positive integrals. Suppose $\{f_n(x)\}$ is a sequence of functions converging in a pointwise monotone manner to 0, $f_n(x) \downarrow 0$.

Since (by the previous theorem) we may write $F = F^+ - F^-$, we have by the definition of F^+: There exists a function $g_n(x)$ such that $0 \leqslant g_n(x) \leqslant f_n(x)$, and $F^+f_n \leqslant Fg_n + 2^{-n}$. Thus, since $F = F - F_m + G_m - H_m$,

$$F^+f_n \leqslant (F - F_m)g_n + (G_m - H_m)g_n + 2^{-n}$$

or

$$F^+f_n \leqslant \|F - F_m\| \, \|f_1\| + G_mf_n + H_mf_n + 2^{-n}.$$

To show that $F^+f_n \to 0$, note that by choosing m large but fixed, $\|F - F_m\| \, \|f_1\|$ can be made arbitrarily small, and then, since G_m and H_m are positive integrals, the remaining terms on the right may be made small at will by letting $n \to \infty$. Thus F^+ is an integral. Similarly F^- is an integral and hence finally $F = F^+ - F^-$ is also an integral. This proves that the integrals form a closed linear manifold.

The Projection Operator. We shall now introduce a bounded linear transformation T whose domain and range is the space $\mathfrak{B}^*$ of linear functionals. T will be defined first for positive linear functionals.

Let $F \in \mathfrak{B}^*$ be a fixed positive functional. Consider an arbitrary positive $f \in \mathfrak{B}$. Let $\{f_n(x)\}$ be an increasing sequence of positive functions in $\mathfrak{B}$ converging pointwise to f, $f_n(x) \uparrow f(x)$. Then $\{Ff_n\}$ is an increasing sequence and since $Ff_n \leqslant Ff$, the sequence $\{Ff_n\}$ has a limit $\geqslant 0$. Now consider the class $\mathfrak{X}$ of all sequences $\{f_n\}$ such that $f_n \uparrow f$ and the class of all limits of the sequences $\{Ff_n\}$. The greatest lower bound of these sequences is a number which depends on F and f and which we denote by Gf. Thus we may write

$$Gf = \text{g.l.b. } (\lim Ff_n), \qquad \{f_n\} \in \mathfrak{X}. \tag{3}$$

We first establish

LEMMA 1. *The functional G is linear and positively homogeneous on the positive functions in $\mathfrak{B}$. That is, if $f(x) \geqslant 0$, $g(x) \geqslant 0$ and $\alpha \geqslant 0$ are given, then $G(f + g) = Gf + Gg$ and $G(\alpha f) = \alpha Gf$.*

Proof. Let $f(x) \geqslant 0$ and $g(x) \geqslant 0$ be given. For a given $\epsilon > 0$ let $\{f_n(x)\}$ be a sequence such that $f_n(x) \uparrow f(x)$ and $\lim Ff_n \leqslant Gf + \epsilon$. Similarly let $\{g_n(x)\}$ be a sequence such that $g_n(x) \uparrow g(x)$ and $\lim Fg_n \leqslant Gg + \epsilon$. Then $\{f_n(x) + g_n(x)\}$ converges to $f(x) + g(x)$ and hence

$$G(f + g) \leqslant \lim F(f_n + g_n) = \lim Ff_n + \lim Fg_n \leqslant Gf + Gg + 2\epsilon.$$

This argument implies that $G(f + g) \leqslant Gf + Gg$.

Now suppose that $\{h_n(x)\}$ is a sequence such that $h_n(x) \uparrow f(x) + g(x)$ and $G(f+g) \geqslant \lim Fh_n - \epsilon$. Let $f_n(x) = f(x) \wedge h_n(x)$ and write $g_n = h_n - f_n$. Then $f_n(x) \uparrow f(x)$ and $g_n(x) \uparrow g(x)$. Thus

$$G(f+g) + \epsilon \geqslant \lim Fh_n = \lim Ff_n + \lim Fg_n \geqslant Gf + Gg.$$

This means that $G(f+g) \geqslant Gf + Gg$. Joining this inequality to the previous one, we have $G(f+g) = Gf + Gg$. Obviously

$$G\alpha f = \alpha(Gf), \qquad \alpha \geqslant 0, \quad f \geqslant 0.$$

The functional G is now extended to the whole of the space $\mathfrak{B}$ in the following manner. We have noted that if $f \in \mathfrak{B}$, then f may be expressed as the difference of two positive functions; for example $f = f^+ - f^-$. Now let f be arbitrary in $\mathfrak{B}$ and suppose $f = g - h$ where $g \geqslant 0$ and $h \geqslant 0$. Define $Gf = Gg - Gh$. The definition is valid, for if $g - h = g' - h'$,

$$Gg + Gh' = G(g + h') = G(g' + h) = Gg' + Gh.$$

LEMMA 2. *The functional G is bounded, positive, and linear over $\mathfrak{B}$. Furthermore*

$$||G|| \leqslant ||F||.$$

Proof. If $f_1, f_2 \in \mathfrak{B}$ write $f_1 = g_1 - h_1, f_2 = g_2 - h_2$ where g_i, h_i are positive. Then

$$\begin{aligned} G(f_1 + f_2) = G([g_1 + g_2] - [h_1 + h_2]) &= G(g_1 + g_2) - G(h_1 + h_2) \\ &= Gg_1 + Gg_2 - Gh_1 - Gh_2 = Gf_1 + Gf_2. \end{aligned}$$

If $\alpha \geqslant 0$, then

$$G\alpha f = G(\alpha[g - h]) = G(\alpha g - \alpha h) = G\alpha g - G\alpha h = \alpha[Gg - Gh] = \alpha Gf.$$

Similarly if $\alpha \leqslant 0$.

In proving that G is bounded, we keep in mind that F is a positive functional. From the definition (3) it is obvious that G is also positive. Thus for $f = f^+ - f^-$ we have

$$Gf = Gf^+ - Gf^- \leqslant Gf^+ \leqslant Ff^+ \leqslant ||F||\,||f^+|| \leqslant ||F||\,||f||.$$

Similarly $Gf \geqslant -||F||\,||f||$. Hence G is bounded and $||G|| \leqslant ||F||$.

Thus we have established a mapping $F \to G$ of the set of bounded positive linear functionals into itself. We shall write $TF = G$. The mapping function T has the properties indicated in

LEMMA 3. *If F, F_1 and F_2 are bounded positive linear functionals and $\alpha \geqslant 0$, then $T(F_1 + F_2) = TF_1 + TF_2$ and $T(\alpha F) = \alpha TF$.*

Proof. Let $F_1 \geqslant 0$, $F_2 \geqslant 0$ and let $f \geqslant 0$. For a given $\epsilon > 0$, let $\{f_n\}$ be such that $f_n \uparrow f$ and

$$\lim (F_1 + F_2)f_n \leqslant T(F_1 + F_2)f + \epsilon.$$

Then

$$T(F_1 + F_2)f + \epsilon \geqslant \lim (F_1 + F_2)f_n = \lim F_1 f_n + \lim F_2 f_n \geqslant TF_1 f + TF_2 f.$$

Now let $\{f_{in}\}$, $i = 1, 2$, be so chosen that $f_{in} \uparrow f$, $TF_i f + \epsilon \geqslant \lim F_i f_{in}$. Let $f_n = f_{1n} \wedge f_{2n}$. Then $f_n \uparrow f$ and $\lim F_i f_n \leqslant \lim F_i f_{in}$. Hence

$$TF_1 f + TF_2 f + 2\epsilon \geqslant \lim F_1 f_{1n} + \lim F_2 f_{2n} \geqslant \lim F_1 f_n + \lim F_2 f_n = \lim (F_1 + F_2) f_n \geqslant T(F_1 + F_2) f.$$

This and the inequality of the preceding paragraphs prove that $T(F_1 + F_2)f = TF_1 f + TF_2 f$. This fact has been established for $f \geqslant 0$. It obviously holds for arbitrary f. Thus $T(F_1 + F_2) = TF_1 + TF_2$. The proof that $T\alpha F = \alpha TF$ runs along similar lines.

We now extend T to all of $\mathfrak{B}^*$. If F is arbitrary in $\mathfrak{B}^*$, then there exist positive functionals G and H such that $F = G - H$. We define: $TF = TG - TH$. By Lemma 3, the definition is admissible.

LEMMA 4. *The transformation T defined above is a bounded and linear transformation of $\mathfrak{B}^*$ into itself. Furthermore if F is a positive integral, $TF = F$.*

Proof. The proof of linearity is straightforward.

We have seen that if F is positive, TF is positive and hence

$$TFf = TF(f^+ - f^-) \leqslant TFf^+ \leqslant Ff^+ \leqslant ||F||\,||f||.$$

Similarly $TFf \geqslant -||F||\,||f||$. Hence $||TF|| \leqslant ||F||$. Now let F be arbitrary and write $F = F^+ - F^-$ where F^+ and F^- are the positive and negative parts of F. Then

$$||TF|| \leqslant ||TF^+|| + ||TF^-|| \leqslant ||F^+|| + ||F^-|| \leqslant 2||F||$$

by the immediately preceding argument.

Finally, if F is a positive integral, then by definition, for any $f \geqslant 0$ and positive sequence $\{f_n\}$ with $f_n \uparrow f$ we have by (2), $Ff_n \to Ff$ hence $TFf = Ff$. If f is arbitrary, the same equation holds, hence $TF = F$.

LEMMA 5. *The transformation T is a projection, that is, $T^2 = T$. The range of the projection consists precisely of all integrals in $\mathfrak{B}^*$.*

Proof. The definition of $G = TF$ for positive F was given in terms of sequences of functions $\{f_n(x)\}$ which converge upward to $f(x)$—see (3). However we may also use series of functions. In fact, let $f(x) \geqslant 0$ be arbitrary and let $\{g_n(x)\}$ be a sequence of positive functions such that $\Sigma g_n(x) = f(x)$. Let $\mathscr{L}$ be the class of all such sequences. Then clearly, we may write (3) in the alternative form

$$(3') \qquad Gf = \text{g.l.b.} \left(\sum_n Fg_n \right), \qquad \{g_n\} \in \mathscr{L}.$$

Now, let $f(x) \geq 0$ and let $F \geq 0$. Let $f = \sum f_n$ where $\{f_n\}$ is any sequence of positive functions. Let $\epsilon > 0$ be given. We find positive functions g_{nm} such that

$$f_n = \sum_m g_{nm}, \quad TFf_n > \sum_m Fg_{nm} - 2^{-n}\epsilon, \qquad n = 1,2,\ldots.$$

Since $TF \geq 0$, $\sum_n TFf_n \leq TFf$ and thus $\sum_n \sum_m Fg_{nm} < \infty$.
More precisely

$$\sum_n TFf_n > \sum_{n,m} Fg_{nm} - \epsilon. \tag{4}$$

Since

$$\sum_{n,m} g_{nm} = f, \quad TFf \leq \sum_{n,m} Fg_{nm}.$$

Substituting in (4) we have $TFf \leq \sum TFf_n + \epsilon$. Since ϵ is arbitrary, this gives $TFf \leq \sum TFf_n$. Since, obviously, $\sum TFf_n \leq TFf$, we have $\sum TFf_n = TFf$.

Now according to (3′)

$$T^2Ff = \text{g.l.b.} \sum_n TFf_n, \qquad \{f_n\} \in \mathfrak{L};$$

thus $T^2Ff = TFf$. The latter identity holds for any $f \geq 0$ and $F \geq 0$ and this leads to the conclusion $T^2 = T$.

To finish the proof of the lemma, it is necessary to show that the range of T is precisely the manifold of integrals. If F is an integral, then by definition $TF = F$. Next suppose that $F \geq 0$ and that $TF = F$. This is precisely the statement that F is an integral. Suppose now that F is arbitrary and that $TF = F$. Writing $F = F^+ - F^-$ we have

$$F = TF = TF^+ - TF^-$$

and since $T^2 = T$,

$$T^2F^+ = TF^+, \quad T^2F^- = TF^-.$$

Now $TF^+ \geq 0$ and $TF^- \geq 0$ and hence both are integrals. Finally, F, which is the difference of two integrals, is an integral. This completes the proof of the lemma.

LEMMA 6. *The transformations T and $I - T$ are positive: that is, $F \geq 0$ implies $TF \geq 0$ and $(I - T)F \geq 0$.*

From the definition of T, it is obvious that $F \geq 0$ and $f \geq 0$ imply $TFf \geq 0$ and $Ff \geq TFf$. This is equivalent to the statement of the lemma.

We now obtain a characterization of the functionals H such that $TH = 0$.

LEMMA 7. *Let $H \geq 0$ and $TH = 0$. Suppose $G \geq 0$ is an integral ($TG = G$) and that $H \geq G$. Then $G = 0$.*

We have $H \geq G \geq 0$. Since T is positive, $0 = TH \geq TG = G \geq 0$, hence $G = 0$.

LEMMA 8. *Let $H \geq 0$ have the property that whenever $G \geq 0$ is such that $TG = G$ and $H \geq G$, then $G = 0$. Then $TH = 0$.*

Since $I - T$ is positive, $(I - T)H \geq 0$, that is $H \geq TH$. However $TH \geq 0$ and $T(TH) = TH$. Thus by the hypothesis concerning G, $TH = 0$.

LEMMA 9. *The projection T is uniquely defined by the properties given in lemmata* 5, 6, 7, 8.

Let T_i $(i = 1, 2)$ be two projections having the indicated properties. Let $\mathfrak{M}_i$ be the set of G such that $T_i G = G$. Similarly let $\mathfrak{N}_i$ be the set of H such that $T_i H = 0$. To prove the lemma, it will be sufficient to show that $\mathfrak{M}_1 = \mathfrak{M}_2$ and $\mathfrak{N}_1 = \mathfrak{N}_2$. The equality $\mathfrak{M}_1 = \mathfrak{M}_2$ is given in Lemma 5.

Before proving that $\mathfrak{N}_1 = \mathfrak{N}_2$ we note that $\mathfrak{N}_i$ is determined by its positive elements. For let F be arbitrary and write $F = F^+ - F^-$. Then the general element in $\mathfrak{N}_i$ is

$$(I - T_i)F = (I - T_i)F^+ - (I - T_i)F^-.$$

Both of these functionals are positive since $I - T_i$ is positive.

Now let $H \geq 0$ be in $\mathfrak{N}_1$. Then $T_1 H = 0$ and hence if G is a positive integral such that $H \geq G$, then $G = 0$ by Lemma 7. Thus by Lemma 8, $T_2 H = 0$, that is $H \in \mathfrak{N}_2$. This shows that $\mathfrak{N}_1 \subset \mathfrak{N}_2$. Since the argument is reversible, $\mathfrak{N}_1 = \mathfrak{N}_2$.

We shall recapitulate these results into our fundamental

THEOREM A. *In the space of linear functionals over the real space $\mathfrak{B}$ there is one and only one projection T with the properties:*

(a) *$TG = G$ if and only if G is an integral.*

(b) *The transformations T and $I - T$ are positive.*

III. COMPLEX SPACES

Up to the present, we have considered real spaces only. We turn to a brief discussion of some details which will show that the theory applies to complex spaces as well. Consider as before a real vector space $\mathfrak{B}_r$ consisting of certain bounded real valued functions defined on an abstract set $\mathcal{E}$. $\mathfrak{B}_r$ is assumed to be algebraically closed under the lattice operations $f \vee g$ and $f \wedge g$ and topologically closed with respect to uniform convergence. Consider now the set $\mathfrak{B}_c$ of all complex-valued functions $f(x) = f_1(x) + i f_2(x)$ where $f_1(x)$, $f_2(x) \in \mathfrak{B}_r$. If we set

$$\|f\| = (\|f_1\|^2 + \|f_2\|^2)^{\frac{1}{2}},$$

it is clear that $\mathfrak{B}_c$ is a complex Banach space. Also $\mathfrak{B}_r$ is a subset of $\mathfrak{B}_c$.

If F is a bounded linear functional on $\mathfrak{B}_r$, it may be extended in a natural way to $\mathfrak{B}_c$. This is done by defining

$$Ff(x) = F(f_1(x) + if_2(x)) = Ff_1(x) + iFf_2(x).$$

The bound of the extended functional F is the same as that of the restricted

F. Functionals F and G of this type may be added and multiplied by complex scalars as follows:

$$(F + G)f = Ff + Gf, \quad [(\alpha + i\beta)F]f = \alpha Ff + i\beta Ff.$$

Thus they constitute a linear manifold contained in $\mathfrak{B}_c^*$. We shall show that this linear manifold coincides with $\mathfrak{B}_c^*$.

To this effect, let F be any bounded linear functional over $\mathfrak{B}_c$. Then F restricted to $\mathfrak{B}_r$ is also linear and bounded. If $f = f(x) \in \mathfrak{B}_r$, let $Ff = F_1 f + iF_2 f$ be the decomposition of Ff into its real and imaginary parts. Then $F_1 f$ and $F_2 f$ are bounded linear functionals over $\mathfrak{B}_r$. Thus, if $F \in \mathfrak{B}_c^*$ and $f \in \mathfrak{B}_r$, there exist $F_1, F_2 \in \mathfrak{B}_r^*$ such that $Ff = F_1 f + iF_2 f$. Now let $f \in \mathfrak{B}_c$, that is, $f = f_1 + if_2$, and let F_1 and F_2 be extended to $\mathfrak{B}_c$. Then it is easy to see that

$$F(f_1 + if_2) = F_1(f_1 + if_2) + iF_2(f_1 + if_2).$$

In other words, $F = F_1 + iF_2$ and each functional on $\mathfrak{B}_c$ is in the linear manifold of the extensions of the functionals on $\mathfrak{B}_r$.

Now consider an arbitrary bounded linear transformation S defined over a real space $\mathfrak{B}_r^*$. We extend S to $\mathfrak{B}_c^*$ by defining

$$S(F_1 + iF_2) = SF_1 + iSF_2.$$

Here F_1 and F_2 denote the extensions to $\mathfrak{B}_c$ of functionals on $\mathfrak{B}_r$. Note that by virtue of the preceding paragraph, every functional $F \in \mathfrak{B}_c^*$ has the form $F = F_1 + i F_2$. Also SF_1 and SF_2 denote the extension to $\mathfrak{B}_c$ of the functionals SF_1 and SF_2 defined originally over $\mathfrak{B}_r$. It may be seen that S is now a bounded linear transformation over $\mathfrak{B}_c^*$.

We apply this procedure to the projection T defined in the preceding pages. By virtue of the linearity of this transformation, it is easy to show that T is a projection over $\mathfrak{B}_c^*$: $T^2 = T$. At this point we extend the notion of an integral. We recall that a real integral F over $\mathfrak{B}_r$ is a linear functional which can be expressed as a difference $F = F_1 - F_2$ where the F_j are positive and such that $f_n \uparrow f$ implies $F_j f_n \to F_j f$, $j = 1, 2$. A complex integral is a linear functional over $\mathfrak{B}_c^*$ of the form $F = G + iH$ where G and H are real integrals. It is now clear that the range of T is the manifold of complex integrals. The other property enunciated for T in Theorem A is obviously valid. We have therefore

THEOREM B. *In the space of linear functionals over the complex space $\mathfrak{B}_c$ there is a unique projection T which has the properties:*

(a) *$TG = G$ if and only if G is an integral.*

(b) *The transformations T and $I - T$ are positive.*

References

1. E. R. Lorch, *Abstract* 250t, Bull. Amer. Math. Soc., *60* (1954), 155.
2. M. A. Woodbury, *Abstract* 167t, Bull. Amer. Math. Soc., *56* (1950), 171.
3. K. Yosida and E. Hewitt, *Finitely additive measures*, Trans. Amer. Math. Soc., *72* (1952), 46–66.

Barnard College
Columbia University
New York

PRODUCTS OF A C-MEASURE AND A LOCALLY INTEGRABLE MAPPING

MARSTON MORSE AND WILLIAM TRANSUE

1. Introduction. Let C be the field of complex numbers and E a locally compact topological space. The authors' theory of C-bimeasures Λ and their Λ-integrals in **(1; 2)** leads to integral representation of bounded operators from A to B' where A and B are MT-spaces as defined in **(3)**. These MT-spaces include the $\mathfrak{L}$-spaces and Orlicz spaces as special cases. The object of this note is to present a theorem on integration necessary in completing the results on MT-spaces. Product measures $g \cdot \mu$, in which g is continuous on E and μ a measure, are introduced by Bourbaki in **(3**, p. 44**)**, and Bourbaki there indicates that products $g \cdot \mu$ in which g is not necessarily continuous will be studied in **(8)**. In this note α is a C-measure and y locally α-integrable in the sense defined below. We draw heavily upon the general theory of integration **(7)**.

Let $\mathfrak{X}$ be the aggregate of relatively compact open subsets X of E. Let ϕ_X be the characteristic function of X. A mapping $y \in \mathrm{C}^E$ will be said to be *locally α-integrable* if for each $X \in \mathfrak{X}$, $y\phi_X$ is α-integrable. The "product" $y \cdot \alpha$ is a C-measure with values

$$\int u \, d(y \cdot \alpha) = \int uy \, d\alpha \qquad [u \in \mathfrak{K}_{\mathrm{C}}(E)]. \tag{1.1}$$

The principal theorem of this note is as follows. Cf. **(3)** for definition of $|\alpha|$.

THEOREM 1.1. *Let α be a* C*-measure on E, and $y \in \mathrm{C}^E$ locally α-integrable. Let $x \in \mathrm{C}^E$ be such that*

$$\int^* |x| \, |y| \, d|\alpha| < \infty, \quad \int^* |x| d|y \cdot \alpha| < \infty. \tag{1.2}$$

Then (i) *these integrals are equal, and* (ii) *if either of the integrals*

$$\int xy \, d\alpha, \quad \int x \, d(y \cdot \alpha) \tag{1.3}$$

exists the other exists and

$$\int xy \, d\alpha = \int x \, d(y \cdot \alpha). \tag{1.4}$$

If the C*-measure α is replaced by a measure μ and y by a locally-integrable $g \in \overline{\mathrm{R}}^E$,* then $g \cdot \mu$ is defined as is $y \cdot \alpha$, replacing $\mathfrak{K}_{\mathrm{C}}$ by $\mathfrak{K}$.

Received January 7, 1957. The work of Dr. Transue on this paper was sponsored by the Office of Ordnance Research, U.S. Army, under contract No. DA-33-019-ORD-1265.

2. A lemma. Let μ be a positive measure on E (sense of Bourbaki). Recall that $\overline{\mathbf{R}}_+$ is the space of positive real numbers completed by the point $+\infty$ and that $\mathfrak{I}_+$ is the space of lower semi-continuous mappings in $\overline{\mathbf{R}}_+^E$.

LEMMA 2.1. *For h and λ in $\overline{\mathbf{R}}_+^E$, with λ locally μ-integrable,*

$$\int^* h\, d(\lambda\cdot\mu) = \int^* h\lambda\, d\mu \tag{2.1}$$

whenever both members of (2.1) *are finite.*

We shall establish the lemma by proving a sequence of statements. We say that (2.1) holds *finitely* if it holds, and if both members are finite.

(a) *If $h \in \mathfrak{I}_+$ is bounded with compact support* (2.1) *holds finitely.*

Set $\lambda\cdot\mu = \beta$. In accord with the definition of $\mu^*(h)$ there exists an increasing sequence (u_n) of $u_n \in \mathfrak{K}_+$ with $u_n \leqslant h$ such that

$$\lim_n \mu(u_n) = \mu^*(h).$$

But h is μ-integrable so that this can be written

$$\lim_n N_1(h - u_n, \mu) = 0.$$

Hence u_n converges to h (p.p. μ), that is, almost everywhere with respect to μ. By the definition of $\lambda\cdot\mu$ and of β^* respectively,

$$\int u_n\lambda\, d\mu = \int u_n\, d\beta \leqslant \int^* h\, d\beta < \infty.$$

Hence by the theorem of Lebesgue,

$$\lim_n \int u_n\lambda\, d\mu = \int h\lambda\, d\mu. \tag{2.2}$$

Let ${}_-\phi_h$ be the set of $u \in \mathfrak{K}_+$ with $u \leqslant h$, filtering for the relation $\leqslant$ (Cf. 3, §6.) From (2.2), the definition of β and of β^*, respectively,

$$\int h\lambda\, d\mu = \sup_{u\in{}_-\phi_h} \int u\lambda\, d\mu = \sup_{u\in{}_-\phi_h} \int u\, d\beta = \int^* h\, d\beta.$$

Since $\int h\lambda\, d\mu = \int^* h\lambda\, d\mu$ (a) follows.

(b) *If h is bounded with compact support*

$$\int^* h\, d\beta \geqslant \int^* h\lambda\, d\mu. \tag{2.3}$$

Let ${}_+\phi_h$ be the set of $q \in \mathfrak{I}_+$ such that $q \geqslant h$, filtering for the relation $\geqslant$. There exists a $q \in {}_+\phi_h$ which is bounded, with compact support. Let S_q be the section of ${}_+\phi_h$ on which $q \geqslant p$. For $p \in S_q$ (a) implies that

$$\int^* p\, d\beta = \int^* p\lambda\, d\mu. \tag{2.4}$$

The definition of $\int^* h\,d\beta$ and (2.4) give

$$\int^* h\,d\beta = \inf_{p \in \mathfrak{Z}_q} \int^* p\,d\beta = \inf_{p \in \mathfrak{Z}_q} \int^* p\lambda\,d\mu \geqslant \int^* h\lambda\,d\mu.$$

(c) *If h is bounded with compact support and μ-integrable,* (2.1) *holds finitely.*

According to (7, p. 151) there exists a decreasing sequence of μ-integrable $p_n \in \mathfrak{Z}_+$ such that for $t \in E$, $p_n(t) \geqslant h(t)$ and $\inf p_n(t) = h(t)$ (p.p. μ). We can suppose each p_n bounded with compact support. Both $p_n\lambda$ and $h\lambda$ are μ-integrable, so that, by the theorem of Lebesgue (noting that the p_n are uniformly bounded),

$$\lim_n \int p_n\lambda\,d\mu = \int h\lambda\,d\mu. \tag{2.5}$$

The definition of $\int^* p\,d\beta$ implies that

$$\int^* p\,d\beta \leqslant \int^* p\lambda\,d\mu. \tag{2.6}$$

Thus

$$\int^* h\,d\beta = \inf_{p \in {}_+\phi_h} \int^* p\,d\beta \leqslant \inf_{p \in {}_+\phi_h} \int^* p\lambda\,d\mu \leqslant \int p_n\lambda\,d\mu.$$

Taken with (2.5) this gives

$$\int^* h\,d\beta \leqslant \int^* h\lambda\,d\mu. \tag{2.7}$$

The inequality is excluded by (2.3), and (c) follows.

(d) *For h bounded with compact support* (2.1) *holds finitely.*

Let K be the compact support of h, and let M be a compact set containing K in its interior. For fixed $\epsilon > 0$ set $\lambda_\epsilon = \lambda + \epsilon\,\phi_M$ and $\beta_\epsilon = \lambda_\epsilon \cdot \mu$. Set $h\lambda_\epsilon = k$ and choose $q \in {}_+\phi_k$ so that the support of q is in M. Let S be the section of the filter ${}_+\phi_k$, of mappings $p \in \mathfrak{Z}_+$ such that $k \leqslant p \leqslant q$. Noting that $\lambda_\epsilon(t) \geqslant \epsilon$ for $t \in M$, and $p(t) = 0$ for $t \in CM$ (the complement of M) and $p \in S$, set

$$h_p(t) = \frac{p(t)}{\lambda_\epsilon(t)},\ t \in M; \quad h_p(t) = 0,\ t \in CM.$$

Then $h_p \geqslant h$, since $\lambda_\epsilon h_p = p \geqslant \lambda_\epsilon h$. That $h_p(t) \geqslant h(t)$ is clear for $t \in M$, and it is trivial when $t \in CM$, since then $h_p(t) = h(t) = 0$. Let $h_p^{(n)}$ be h_p truncated at the level n. From (c)

$$\int^* h_p^{(n)}\,d\beta_\epsilon = \int^* h_p^{(n)}\lambda_\epsilon\,d\mu.$$

On letting $n \uparrow \infty$ it follows from (7, p. 111) that

$$\int^* h_p\,d\beta_\epsilon = \int^* h_p\,\lambda_\epsilon\,d\mu = \int^* p\,d\mu.$$

Since $h \leqslant h_\epsilon$ and $\beta \leqslant \beta_\epsilon$ this gives

$$\int^* h\, d\beta \leqslant \inf_{p \in S} \int^* p\, d\mu = \int^* h\lambda_\epsilon\, d\mu. \tag{2.8}$$

Now $\epsilon > 0$ is arbitrary in (2.8) so that

$$\int^* h\, d\beta \leqslant \int^* h\lambda\, d\mu < \infty.$$

The inequality is excluded by (2.3), and (d) follows.

(e) *If h has compact support* (2.1) *holds.*

For each integer $n > 0$, let $h^{(n)}$ be h, truncated at the level n. Then by (d)

$$\int^* h^{(n)}\, d\beta = \int^* h^{(n)}\lambda\, d\mu. \tag{2.9}$$

As $n \uparrow \infty$, $h^{(n)}(t)$ converges increasing to $h(t)$. It follows from (7, p. 111) that (2.1) holds.

(f) *If a member of* (2.1) *is finite the other is at least as great.*

Suppose the right member of (2.1) is finite. It then follows from (7, Lemma 2, p. 194) that E is a disjoint union $H \cup M$ in which $h\lambda\, \phi_M$ is μ-negligible and H is the increasing union of a sequence of compact sets K_n. Set

$$h\phi_{K_n} = h_n,\ h\phi_H = h',\ h\phi_M = h''.$$

By (e),

$$\int^* h_n\, d\beta = \int^* h_n\lambda\, d\mu.$$

Now $h_n(t)$, increasing, converges to $h'(t)$ as $n \uparrow \infty$, so that it follows as in the proof of (e) that

$$\int^* h'\, d\beta = \int^* h'\lambda\, d\mu. \tag{2.10}$$

Since $h = h' + h''$ and $\lambda h''$ is μ-negligible,

$$\mu^*(h\lambda) \leqslant \mu^*(h'\lambda) + \mu^*(h''\lambda) = \mu^*(h'\lambda) \leqslant \mu^*(h\lambda) \tag{2.11}$$

implying equalities throughout (2.11). Hence

$$\int^* h\, d\beta \geqslant \int^* h'\, d\beta = \int^* h'\lambda\, d\mu = \int^* h\lambda\, d\mu$$

so that (f) follows when $\int^* h\lambda\, d\mu < \infty$.

The proof of (f) when $\int^* h\, d\beta < \infty$ is similar, so that (f) is true.

Lemma 2.1 follows from (f).

Note. Since this paper reached the hands of the Editors, (8) has appeared. Our Lemma 2.1 should be compared with B. Prop. 2, p. 43, noting that

Bourbaki uses $\overline{\int}^*$ while we use $\int^*$. Prop. 2 of Bourbaki follows at once from (e) of this section. Conversely Prop. 2 of Bourbaki implies (e), and one may then continue, as we have done, with a proof of our Lemma 2.1.

One should also compare our Th. 1.1 with **(7**, Th. 1 p. 43**)**, noting that Bourbaki is concerned here with "essential" integrals of mappings into a Banach space with respect to positive measures, whereas we are concerned with integrals of mappings into C with respect to C-measures and product C-measures $y \cdot \alpha$. Our theorem depends on the fundamental formula $|y \cdot \alpha| = |y| \cdot |\alpha|$ of §3 and the lemmas of §4. The deeper connections between our theorem and that of Bourbaki will be brought out in a note **(5)**, presently to appear. The reader may also note the connection with the Radon-Nikodym Theorem as shown by Bourbaki.

3. The formula $|y \cdot \alpha| = |y| \cdot |\alpha|$. This formula is equivalent to the formula

$$\int f\,|y|\,d|\alpha| = \int f\,d|y \cdot \alpha| \qquad [f \in \mathfrak{R}_+]. \tag{3.1}$$

Set $\beta = y \cdot \alpha$. It follows from (1.1) and the definition of $|\beta|$ in **(3**, (3.3)**)** that

$$|\beta|(f) \leqslant \int f\,|y|\,d|\alpha|, \qquad [f \in \mathfrak{R}_+]. \tag{3.2}$$

It remains to show that the inequality is excluded. We shall need the following.

(i) *Let H be the set of points t at which $y(t) \neq 0$. The function σ with values*

$$\sigma(t) = \frac{y(t)}{|y(t)|}, \qquad t \in H; \sigma(t) = 0, t \in CH$$

is $|\alpha|$-measurable.

Since y is measurable, $|y|$ is measurable. The function τ with values $\tau(t) = |y(t)|^{-1}$ for $t \in H$, and $\tau(t) = 0$ when $t \in CH$, is measurable, as one sees with the aid of **(7**, Prop. 9, p. 192**)**. Since $\sigma = \tau y$, σ is also measurable. Thus (i) holds.

To show that the inequality is excluded in (3.2) we can suppose without loss of generality that $\max f(t) = 1$. Let K be the compact support of f. Let $\epsilon > 0$ be arbitrary. Set $\phi_K|y| = \lambda$ and $|\alpha| = \mu$. Since λ is μ-integrable there exists a $\lambda_0 \in \mathfrak{R}_+$ such that

$$\int |\lambda - \lambda_0|\,d\mu < \epsilon. \tag{3.3}$$

For our purposes a λ_0 is needed which is positive on K. To this end λ_0 is modified as follows. Let K_1 be compact and such that for some open set U, $K_1 \supset U \supset K$. Let $\lambda_1 \in \mathfrak{R}_+$ map E into $[0, 1]$ with support K_1 and with $\lambda_1(t) = \phi_K(t)$ for $t \in K$. If $c > 0$ is sufficiently small, and if one sets $\lambda_2 = \lambda_0 + c\lambda_1$ then

$$\int |\lambda - \lambda_2|\, d\mu < \epsilon \tag{3.4}$$

while $\lambda_2(t) \geqslant c > 0$ for $t \in K$. From (3.4) and the assumption that $\max f(t) = 1$ it follows that

$$\int f\lambda\, d\mu \leqslant \int f\lambda_2\, d\mu + \epsilon. \tag{3.5}$$

Now $f\lambda_2$ is in $\mathfrak{R}_+$. In accord with the definition of $|\alpha|$, and with u and $v \in \mathfrak{R}_C$

$$\int f\lambda_2\, d\mu = \sup_{|u| < f\lambda_2} \left| \int u\, d\alpha \right| = \sup_{|v| < f} \left| \int v\lambda_2\, d\alpha \right|. \tag{3.6}'$$

It follows from (3.4) that

$$\left| \int v\lambda_2\, d\alpha \right| \leqslant \left| \int v\lambda\, d\alpha \right| + \left| \int v(\lambda_2 - \lambda) d\alpha \right| \leqslant \left| \int v\lambda\, d\alpha \right| + \epsilon \tag{3.6}''$$

for $|v| \leqslant f$. From (3.5) and (3.6), for some $v \in \mathfrak{R}_C$ with $|v| \leqslant f$,

$$\int f\lambda\, d\mu \leqslant \left| \int v\lambda\, d\alpha \right| + 3\epsilon = \left| \int v\bar{\sigma}\, y\, d\alpha \right| + 3\epsilon \tag{3.7}$$

introducing σ as defined above.

Now $\bar{\sigma}$, as the conjugate of σ, is μ-measurable. There accordingly exists (**7**, Prop. 1, p. 180) a compact set $H \subset K$ such that $\bar{\sigma}|H$ is continuous and the μ-measure, say ω, of $K \cap CH = M$, is arbitrarily small. In particular one can suppose ω so small that $\int \phi_M f|y|\, d\mu < \epsilon$. Let g denote a continuous extension of $\bar{\sigma}|H$ to K with $\max|g(t)| < 1$ (see Note). Since $(\bar{\sigma} - g)|H = 0$ and $|vg| \leqslant |v| \leqslant f$

$$\left| \int vy(\bar{\sigma} - g) d\alpha \right| = \left| \int \phi_M vy(\bar{\sigma} - g) d\alpha \right| \leqslant 2 \int \phi_M f|y|\, d\mu \leqslant 2\epsilon$$

so that with $u \in \mathfrak{R}_C$

$$\left| \int vy\, \bar{\sigma}\, d\alpha \right| - 2\epsilon \leqslant \left| \int vgy\, d\alpha \right| \leqslant \sup_{|u| < f} \left| \int uy\, d\alpha \right| = |\beta|(f). \tag{3.8}$$

From (3.7) and (3.8)

$$\int f\lambda\, d\mu \leqslant |\beta|(f) + 5\epsilon.$$

Hence the inequality must be excluded in (3.2) and the relation $|y \cdot \alpha| = |y| \cdot |\alpha|$ follows.

The "product" $y \cdot \alpha$ is clearly doubly distributive. In the case of a real measure μ and real g, $|g \cdot \mu| = |g| \cdot |\mu|$, as the above proof shows after trivial notational modifications. It follows immediately that

$$(g \cdot \mu)^+ = g^+ \cdot \mu^+ + g^- \cdot \mu^-, \quad (g \cdot \mu)^- = g^- \cdot \mu^+ + g^+ \cdot \mu^-$$

Note. It follows from a theorem of Urysohn **(6**, p. 62**)** that a continuous extension f of $\tilde{\sigma}|H$ over K exists. To obtain from f a continuous extension g of $\tilde{\sigma}|H$ such that $\max|g(t)| < 1$, set $g(t) = f(t)$ when $|f(t)| \leqslant 1$, and at all other points of K set $g(t) = f(t)/|f(t)|$.

4. Proof of Theorem 1.1. We need a lemma on measurability. Recall that y is locally α-integrable.

LEMMA 4.1. *If*

$$\int^* |x||y|d|\alpha| = \int^* |x|d|y\cdot\alpha| < \infty \tag{4.1}$$

then (i) *for an arbitrary subset M of E then α-negligibility of $\phi_M xy$ is equivalent to the $(y\cdot\alpha)$-negligibility of $\phi_M x$, and* (ii) *the α-measurability of xy is equivalent to the $(y\cdot\alpha)$-measurability of x.*

One must first verify the fact that (4.1) has the form of the relation (2.1) if one sets $\mu = |\alpha|$ and $|y| = \lambda$. This follows from the formula $|y\cdot\alpha| = |\alpha|\cdot|y|$ just established. One can accordingly apply Lemma 2.1 as follows. For an arbitrary subset M of E

$$\int^* \phi_M|x||y|d|\alpha| = \int^* \phi_M|x|d|y\cdot\alpha| < \infty. \tag{4.2}$$

For the two members of (4.2) are finite and hence equal by Lemma 2.1. Statement (i) follows.

(α) *Set $\beta = y\cdot\alpha$. If x is β-measurable, xy is α-measurable.*

Let K be compact. According to the Bourbaki definition of measurability K is a disjoint union $K = H \cup M$, where H is a countable union of compact sets K_n on each of which x is continuous, and where M is β-negligible. Then $\phi_M x$ is β-negligible, and so by (i), $\phi_M xy$ is α-negligible and hence α-measurable **(7**, Prop. 6, p. 184**)**. Further $\phi_H x$ is α-measurable and hence $\phi_H xy$. Finally $\phi_K xy = \phi_H xy + \phi_M xy$ is α-measurable, and hence xy. **(**"Principal of localization", **7**, p. 182.**)**

(β) *If x is α-measurable, x is β-measurable.* Let K, H, M be chosen as in the proof of (α) except that M shall here be α-negligible. Then $\phi_M x$ is β-negligible by (i), and hence β-measurable. Thus x is β-measurable since $\phi_K x = \phi_H x + \phi_M x$.

(γ) *If xy is α-measurable, x is β-measurable.* Let M be the subset of E on which $y(t) \neq 0$. Then M is α-measurable. Set $H = CM$. The relation

$$\phi_M x = \frac{\phi_M xy}{\phi_H + y} \tag{4.3}$$

shows that $\phi_M x$ is α-measurable, since both numerator and non-vanishing denominator are α-measurable. One can apply (β), replacing x by $\phi_M x$, since (4.1) holds with $\phi_M x$ replacing x. Hence $\phi_M x$ is β-measurable. Now $\phi_H xy$

vanishes, so that by (i) $\phi_H x$ is β-negligible and hence β-measurable. Hence $x = \phi_M x + \phi_H x$ is β-measurable.

Lemma 4.1 (ii) follows from (α) and (γ).

Statement (ii) is conditioned in Lemma 4.1 by (4.1). Actually this condition can be dropped.

COROLLARY 4.1. *For a locally α-integrable y, the α-measurability of xy is equivalent to the $(y\cdot\alpha)$-measurability of x.*

We commence with the following.

(a) *If xy is α-measurable $x^{(n)}y$ is α-measurable.* It is understood that $x^{(n)} = x_1^{(n)} + ix_2^{(n)}$ where $x_1 + ix_2$ is a Gaussian decomposition of x. If M is taken as in (γ) then $\phi_M x$ is α-measurable. Hence $\phi_M x^{(n)}$ and consequently $(\phi_M x^{(n)})y = x^{(n)}y$ is α-measurable.

(b) *If $x^{(n)}y$ is α-measurable for each positive integer $n > 0$, xy is α-measurable.* For $x^{(n)}y$ converges pointwise to xy as $n \uparrow \infty$.

From (a) and (b) we conclude that the Corollary is true if true for bounded x. We therefore prove the following.

(c) *The Corollary is true for bounded x.*

Let K be compact and set $z = \phi_K x$. Now z is bounded with compact support. Hence (cf. (d) of §2),

$$\int^* |z||y|d|\alpha| = \int^* |z|d|(y\cdot\alpha)|$$

and Lemma 4.1 applies, so that the Corollary is true if z replaces x. By the principle of the localization of measurability, (c) is true as stated.

The Corollary follows as indicated above.

Proof of Theorem 1.1. Statement (i) of the theorem follows from Lemma 2.1 since $|y\cdot\alpha| = |y|\cdot|\alpha|$ so that one can identify $|\alpha|$ with μ in Lemma 2.1. *Assuming then that* (4.1) *holds* we prove the following. Set $y\cdot\alpha = \beta$.

(a) *If x is β-integrable and if (x_n) is a sequence of β-integrable mappings such that $|x_n| \leqslant |x|$ and $x_n(t)$ converges to $x(t)$* (p.p. β), *then* (1.4) *holds provided*

$$\int x_n\, d\beta = \int x_n y\, d\alpha \qquad (n = 1,2,\dots). \tag{4.4}$$

Let H be the set of points t on which $x_n(t)$ converges to $x(t)$. Set $M = CH$. Then M is β-negligible so that $\phi_M xy$ is α-negligible (Lemma 4.1). Since $|x_n| \leqslant |x|$, $\phi_M x_n y$ is α-negligible. Now $\phi_H x_n y$ converges pointwise to $\phi_H xy$. Moreover

$$|\phi_H x_n y| \leqslant \phi_H |x||y|.$$

Since $\alpha^*(\phi_H|x||y|) < \infty$ by hypothesis, the theorem of Lebesgue and the α-negligibility of $\phi_M xy$ imply that (7, p. 140, Th. 6)

$$\lim_{n\uparrow\infty} \int \phi_H x_n y d\alpha = \int \phi_H xy\, d\alpha = \int xy\, d\alpha.$$

Thus

$$\int x\,d\beta = \lim_{n\uparrow\infty}\int x_n\,d\beta = \lim_{n\uparrow\infty}\int x_n y\,d\alpha \qquad \text{from (4.4)}$$

$$= \lim_{n\uparrow\infty}\int \phi_H x_n y d\alpha = \int xy\,d\alpha.$$

(b) *If* $x \in \mathfrak{F}_+$ *is* β*-integrable then* (1.4) *holds.* For one can satisfy the conditions on (x_n) in (a) by proper choice of $x_n \in \mathfrak{R}_+$. The relation (4.4) holds by definition of β.

(c) *If* $x \in \overline{\mathrm{R}}_+^B$ *is bounded and upper semi-continuous with compact support, then* (1.4) *holds.* Let K be the compact support of x, and let $u \in \mathfrak{R}_+$ be such that $u(t) \geqslant x(t)$ on K. Then $u - x \in \mathfrak{F}_+$ and is β-integrable. From (b), (1.4) holds with $u - x$ replacing x. Hence (1.4) holds.

(d) *If* $x \in \overline{\mathrm{R}}_+^B$ *is* β*-integrable and upper semi-continuous with compact support, then* (1.4) *holds.* The truncation $x^{(n)}$ satisfies the conditions on x_n of (a), as follows from (c). Hence, from (a), (1.4) holds.

(e) *If* $x \in \overline{\mathrm{R}}_+^B$ *is* β*-integrable, then* (1.4) *holds.* For one can satisfy the conditions on (x_n) in (a), possibly excepting (4.4) by choice of x_n which are upper semi-continuous with compact support **(7**, p. 151**)**. By virtue of (d), (4.4) will then hold, and (1.4) follows.

(f) *If* x *is* β*-integrable* (1.4) *holds.* Let h be any one of the Riesz components **(3**, §4**)** of x. Since (4.1) holds by hypothesis, (4.1) holds, h replacing x, since both members of (4.1) are then finite and hence equal by Lemma 2.1. Now h is β-integrable **(3**, Cor. 9.2**)**. By (e), (1.4) holds, h replacing x. Hence (1.4) holds for x.

(g) *If* xy *is* α*-integrable* (1.4) *holds.* By Lemma 4.1, x is β-measurable, and since $\beta^*(x) < \infty$ by hypothesis, x is β-integrable. Hence (1.4) holds by (f).

Theorem 1.1 (ii) follows from (f) and (g).

5. The relation 2.1. We here present lemmas which facilitate the application of Theorem 1.1. We shall term a countable union of sets $X_n \in \mathfrak{X}$ an ω-set (cf. §1 for definition of $\mathfrak{X}$). We return to λ, h, μ of Lemma 2.1 and refer to the conditions

$$\int^* h\,d(\lambda\cdot\mu) \geqslant \int^* h\lambda\,d\mu, \tag{5.1}$$

$$\int^* h\,d(\lambda\cdot\mu) \leqslant \int^* h\lambda\,d\mu. \tag{5.2}$$

LEMMA 5.1. *If* $\lambda = \lambda' + \lambda''$, *where* $\lambda' \geqslant 0$, $\lambda'' \geqslant 0$, λ'' *is* μ*-negligible, and the support of* λ' *is contained in an* ω*-set* M, *then* (2.1) *holds.*

We shall need the following.

(a) *If ν is a positive measure with support H, then for $g \in \overline{\mathfrak{R}}_+^*$*

$$\int^* g\, d\nu = \int^* g\phi_H\, d\nu. \tag{5.3}$$

To verify this relation let $p \in \mathfrak{I}_+$ be such that $p\phi_H = 0$ and $p(t) = +\infty$ when $t \in CH$. Let $g\phi_H = k$. Then $p \geqslant g - k$ so that $\nu^*(p) \geqslant \nu^*(g - k)$. However,

$$\nu^*(p) = \sup_{f \leqslant p} \nu(f) \qquad f \in \mathfrak{L}_+.$$

Since such $f \leqslant p$ vanish on H, and since H is the support of ν, $\nu^*(f) = 0$. Hence $0 = \nu^*(p) = \nu^*(g - k)$. Relation (5.3) follows.

To establish the lemma suppose first that $\lambda'' = 0$, so that $\lambda = \lambda'$. Suppose further that M is the countable union of increasing sets $X_n \in \mathfrak{X}$. Set

$$h_n = h\phi_{X_n}.$$

Since h_n has a compact support

$$\int^* h_n d(\lambda \cdot \mu) = \int^* h_n \lambda\, d\mu \qquad \text{by (e) of §2.} \tag{5.4}$$

Since $h_n(t)$ converges increasing to $\phi_M h(t)$ as $n \uparrow \infty$, it follows from (**7**, p. 111 Corollary) that

$$\int^* \phi_M h\, d(\lambda \cdot \mu) = \int^* \phi_M h\lambda\, d\mu = \int^* h\lambda\, d\mu. \tag{5.5}$$

The support H of $\lambda \cdot \mu$ is contained in the support of λ. Hence $H \subset M$. Two applications of (5.3) then give

$$\int^* h\, d(\lambda \cdot \mu) = \int^* \phi_H h\, d(\lambda \cdot \mu) = \int^* \phi_M h\, d(\lambda \cdot \mu). \tag{5.6}$$

Relation (2.1) follows from (5.5) and (5.6).

In the case of the general $\lambda = \lambda' + \lambda''$

$$\int^* h\, d(\lambda' \cdot \mu) = \int^* h\lambda'\, d\mu \tag{5.7}$$

as we have just seen. But $\lambda \cdot \mu = \lambda' \cdot \mu$ and $h\lambda''$ is μ-negligible. Hence (5.7) implies (2.1).

A mapping function which vanishes in the complement of an ω-set M will be termed an *ω-function with ω-set M.*

LEMMA 5.2 (i). *If $h\lambda$ is μ-equivalent to an ω-function with ω-set M then* (5.1) *holds.*

(ii) *If h is $(\lambda \cdot \mu)$-equivalent to an ω-function with ω-set M then* (5.2) *holds.*

Proof of (i). Let $\lambda h = \lambda h' + h''$ where h' is the ω-function $h\phi_M$ with ω-set M, and h'' is μ-negligible. Suppose that M is the countable union of sets $X_n \in \mathfrak{X}$ and set $h_n = h\,\phi_{X_n}$. Then h_n converges to $\phi_M h$ so that

$$\int^* \phi_M h \, d(\lambda \cdot \mu) = \int^* \phi_M h \, \lambda \, d\mu = \int^* h\lambda \, d\mu$$

as in the proof of (5.5). Relation (5.1) follows.

The proof of (ii) is similar.

Lemma 5.2 (i) applies to a product λh in $\mathfrak{L}^p(\mu)$, since such a λh is μ-equivalent to an ω-function by **(7**, Lemma 2, p. 194**)**. An $h \in L^p(\lambda \cdot \mu)$ is similarly relevant to Lemma 5.2 (ii).

A space E which is "countable at infinity" is an ω-set by definition, so that on such a space (2.1) holds by virtue of Lemma 5.1.

If μ is a bounded measure, E is the union of an ω-set and a μ-negligible set, so that each mapping in C^E is μ-equivalent to an ω-function. A similar remark applies to a bounded measure $\lambda \cdot \mu$.

There are many other combinations of special conditions relevant to these lemmas.

6. Two counter-examples. The fact that inequalities appear in (i) and (ii) of Lemma 5.2 raises the question as to whether or not there are examples in which the equality is excluded. The question also arises in connection with Lemma 2.1. The following two examples show that the inequality may occur in either (i) or (ii).

Example 6.1. Let $H \subset E$ be locally μ-negligible, but not μ-negligible. Such a μ and set exist **(7**, p. 184**)**. Set $h = \phi_H$, $\lambda = \phi_{CH}$. Then h and λ are bounded and μ-measurable, so that λ is locally μ-integrable. Moreover h is not μ-integrable since it is not μ-negligible **(7**, p. 195**)**. Hence $\mu^*(h) = \infty$. For $u \in \mathfrak{L}_+$, hu is μ-negligible so that

$$\int u \, d\mu = \int hu \, d\mu + \int \lambda u \, d\mu = \int \lambda u \, d\mu.$$

It follows that $\mu = \lambda \cdot \mu$. Since $h\lambda = 0$, $\mu^*(h\lambda) = 0$, while

$$(\lambda \cdot \mu)^*(h) = \mu^*(h) = \infty.$$

Thus $\mu^*(h\lambda) < (\lambda \cdot \mu)^*(h)$. Note finally that $h\lambda$ is μ-equivalent to an ω-function, the null function.

Example 6.2. Take H as in Example 6.1. Set $\lambda = \phi_H$ and $h = \phi_E$. Then $\lambda \cdot \mu$ is a null measure so that $(\lambda \cdot \mu)^*(h) = 0$. Moreover $\mu^*(h\lambda) = \mu^*(\lambda) = \infty$. Note also that h is $(\lambda \cdot \mu)$-equivalent to a null function.

References

1. M. Morse and W. Transue, C-*bimeasures* A *and their superior integrals*, Rend. Circ. Mat. Palermo, *4* (1955), 270–300.
2. ———, C-*bimeasures* A *and their integral extensions*, Ann. Maths., *64* (1956), 480–504.
3. ———, *Semi-normed vector spaces with duals of integral type*, J. d'Analyse Math., *4* (1955), 149–186.
4. ———, *Vector subspaces A of* C^B *with duals of integral type*, J. Math. Pures Appl., *00* (1957), to be published.
5. ———, *A comparison of two theorems on integration*, Proc. Nat. Acad. Sci., *00* (1957), to be published.
6. N. Bourbaki, *Éléments de Mathématique, 8, Topologie Générale*, Chap. IX (Paris, 1948).
7. ———, *Éléments de Mathématique, 13, Intégration*, Chap. I-IV (Paris, 1952).
8. ———, *Éléments de Mathématique, 21, Intégration*, Chap. V (Paris, 1956)

Institute for Advanced Study
Princeton, New Jersey

Kenyon College
Gambier, Ohio

A TAUBERIAN THEOREM FOR THE RIEMANN-LIOUVILLE INTEGRAL OF INTEGER ORDER

C. T. RAJAGOPAL

1. Notation. Let $s(x)$ be a function integrable[1] in every finite interval of $x > 0$. Then the Riemann-Liouville integral of $s(x)$, of order $\alpha > 0$, is defined for $x > 0$ by

$$s_\alpha(x) = \frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} s(t)\,dt. \tag{1}$$

The object of this note is to prove a Tauberian theorem for $s_\alpha(x)$ in the case in which α is a positive integer p, employing certain difference formulae due to Karamata **(4**, Lemma 2**)** and Bosanquet **(1**, Theorem 1**)** used already for a broadly similar purpose in an earlier paper **(12)** where α is any positive number.

Adopting a familiar notation, we shall write

$$C_\alpha(x) = \frac{\Gamma(\alpha+1)\, s_\alpha(x)}{x^\alpha}, \qquad \alpha > 0, \tag{2}$$

$$C_0(x) = s_0(x) \equiv s(x),$$

and say that $s(x)$ is summable by the Cesàro mean of order α, or briefly, summable (C, α), to sum l, when

$$\lim_{x\to\infty} C_\alpha(x) = l,$$

l denoting a finite number as everywhere in this note. When $\lim C_\alpha(x)$ does not exist, as in the principal results of this note, it is convenient to write

$$\liminf_{x\to\infty} C_\alpha(x) = \underline{C}_\alpha, \qquad \limsup_{x\to\infty} C_\alpha(x) = \bar{C}_\alpha. \tag{3}$$

2. Scope of the main result. The following theorems, stated in the notation explained above, are known, at least in some part or form; and all of them turn out to be easy consequences or modifications of the single main result of this note featured as Theorem I.

THEOREM A. *If $s(x)$ is an integral,*

$$s'(x) = O_R(x^{q-p-1}) \text{ as } x \to \infty \tag{4}$$

for almost all $x > 0$, p and q being real numbers of which the former is a positive integer, then

$$\frac{s_p(x)}{x^q} \to l \qquad (x \to \infty) \tag{5}$$

Received October 3, 1956.

[1]In this note integrability and integrals are always in the sense of Lebesgue.

implies

$$\frac{s(x)}{x^{q-p}} \to lq(q-1)\dots(q-p+1). \tag{6}$$

Theorem A was first proved by Doetsch **(2**, p. 174, Theorem II**)** with the restriction $q > p+1$[2] which was subsequently removed by Obrechkoff **(5)**. A special case of Theorem A with $p = 1$ had been proved earlier by Hardy and Littlewood **(13**, p. 194, Corollary 4.4*a***)**, while a more general form of the theorem, with the positive integer p replaced by any positive number α was obtained later by Parthasarathy and Rajagopal **(6**, Theorem B, Case (2)**)**.

A generalization of Theorem A is the following theorem wherein (4) is replaced by (4′), a condition which evidently holds whenever (4) holds.

THEOREM A′. *If $s(x)$ is such that*

$$\lim_{\lambda\to 1+0} \limsup_{t\to\infty} \sup_{t\le t'\le \lambda t} \frac{s(t')-s(t)}{t^{q-p}} < 0, \tag{4′}$$

then (5) *implies* (6).

Theorem A′ and, in fact, its extension when the limit in (5) does not exist, are both included in the main result of this note whose proof is by the method used by Parthasarathy and Rajagopal **(6)** to obtain the extension of Theorem A in which p is replaced by any $\alpha > 0$.

The case $q = p$ of Theorem A′ is the classical result stated next.

THEOREM B. *If $s(x)$ is slowly increasing, that is,*

$$\lim_{\lambda\to 1+0} \limsup_{t\to\infty} \sup_{t\le t'\le \lambda t} \{s(t')-s(t)\} < 0,$$

and summable (C, p) *to l,*[3] *then $s(x)$ converges to l as $x \to \infty$.*

The following theorem is a companion to Theorem B; its case $p = 1$ has been proved in a somewhat different form by Pitt **(7)**.

THEOREM C. *In Theorem* B *the condition of slow increase of $s(x)$ can be replaced by the following condition, without any other change:*

$$\lim_{\lambda\to 1+0} \limsup_{t\to\infty} \sup_{t\le t'\le \lambda t} \left| \frac{1}{(\lambda-1)t} \int_t^{t'} \{s(u)-s(t)\}du \right| = 0. \tag{7}$$

A classical particularization of Theorem C is that in which (7) is replaced by the condition of slow oscillation of $s(x)$ which clearly implies (7).[4] A

[2]The case $q = p+1$ of Theorem A, with $s'(x)$ replaced by $s(x)$, gives the well-known theorem: *if $s(x)$ is bounded on one side and summable* (C, $p+1$) *to l, then it is summable* (C, 1) *to l.*

[3]In virtue of the first theorem of consistency for Cesàro summability, p in such cases may be replaced by any $\alpha > 0$.

[4]A condition which is effectively the same as that of slow oscillation is the "high-indices" condition, $\liminf \lambda_{n+1}/\lambda_n > 1$, when $s(x)$ is the λ_n − step function defined in the concluding remarks.

simple modification of the case $p = 1$ of Theorem C is like Pitt's theorem (7, Theorem 1) and unlike any of the classical Tauberian theorems for Cesàro summability in having no exact counterpart for Abel summability, that is, in not being always true when Cesàro summability is replaced by Abel summability (without any other change).

The last theorem to be now given includes Theorem B in the case $p = 1$.

THEOREM D. *If*

$$\limsup_{t\to\infty} \sup_{1\le t'\le t} \{s(t') - s(t)\} = w_1(\lambda), \tag{8}$$

then, for $0 < \theta < 1 < \lambda$,

$$-(\lambda - 1)\,\underline{C}_0 \leqslant -\lambda \underline{C}_1 + \bar{C}_1 + \int_1^\lambda w_1(t)dt,$$

$$(1 - \theta)\,\bar{C}_0 \leqslant -\theta\,\underline{C}_1 + \bar{C}_1 + \int_\theta^1 w_1(t)dt.$$

Theorem D is Karamata's[5] (3, Satz 1, first part), and its significance lies in the fact that it includes certain best-possible inequalities connecting C_0 and $\bar{C}_0$ with $\underline{C}_1$ and $\bar{C}_1$ first obtained by Fekete and Winn under the condition (8) with $w_1(\lambda) \leqslant K \log \lambda$. A generalization of Theorem D proved by me elsewhere (8, Lemma 3) is in the form of inequalities connecting $\underline{C}_\alpha$ and $\bar{C}_\alpha$ with $\underline{C}_{\alpha+1}$ and $\bar{C}_{\alpha+1}$ under condition (8) again. On the other hand, the generalization of Theorem D in this note, viz. Theorem I, takes the form of inequalities connecting lim inf $s(x)/x^{q-p}$ and lim sup $s(x)/x^{q-p}$ with lim inf $s_p(x)/x^q$ and lim sup $s_p(x)/x^q$ under a Tauberian condition which reduces to (8) when $q = p$. When

$$\liminf s_p(x)/x^q = \limsup s_p(x)/x^q,$$

the inequalities of Theorem I lead, in a special case, to Corollary I(1) which is Theorem A′ in the notation of Theorem I. When $q = p$, the inequalities of Theorem I become the inequalities of Corollary I(2) connecting $\underline{C}_0$ and $\bar{C}_0$ with $\underline{C}_p$ and $\bar{C}_p$, the case $p = 1$ of the latter inequalities constituting Theorem D. Modifications of the aforesaid inequalities connecting $\underline{C}_0$ and $\bar{C}_0$ with $\underline{C}_p$ and $\bar{C}_p$ are obtained in Corollary I(3) when (8) is replaced by the following condition implicit in (7):

$$\limsup_{t\to\infty} \sup_{1\le t'\le t} \left| \frac{1}{t(\lambda - 1)} \int_t^{t'} \{s(u) - s(t)\}du \right| = \Omega(\lambda).$$

In brief, Corollary I(2) and Corollary I(3) extend Theorem B and Theorem C respectively on the lines of Theorem D. Corollary I(4) following them refashions the case $p = 1$ of Corollary I(3) so as to produce in particular the (C, 1) summability theorem mentioned earlier as having no counterpart for Abel summability.

[5]Karamata's theorem has been restated here to match Theorem I, with $-s(x)$ in place of his $s(x)$.

3. The main result. The statement of this result, appearing as Theorem I, is necessarily elaborate by reason of the comprehensive character of the theorem. But the proof of the theorem is in essentials as simple as that of Theorem D, requiring nothing more than the Karamata-Bosanquet difference formulae referred to at the outset and embodied in the following lemmas easily verifiable by induction.

LEMMA 1. *If* $h > 0$, $p = 1,2,3, \ldots$, *then*

$$\Delta_h^p s_p(x) \equiv \sum_{\nu=0}^{p} (-1)^\nu \binom{p}{\nu} s_p(x + \overline{p-\nu}h)$$
$$= \int_x^{x+h} dt_1 \int_{t_1}^{t_1+h} dt_2 \ldots \int_{t_{p-1}}^{t_{p-1}+h} s(t)dt.$$

LEMMA 2. *If* $k > 0$, $p = 1,2,3, \ldots$, *then*

$$\Delta_{-k}^p s_p(x) \equiv \sum_{\nu=0}^{p} (-1)^\nu \binom{p}{\nu} s_p(x - \nu h)$$
$$= \int_{x-k}^{x} dt_1 \int_{t_1-k}^{t_1} dt_2 \ldots \int_{t_{p-1}-k}^{t_{p-1}} s(t)dt.$$

THEOREM I. *Let* $s(x)$, *integrable in every finite interval of* $x \geqslant 0$, *be such that, for* $\lambda > 1$, *one of the following two conditions holds and consequently the other also:*

$$\limsup_{t\to\infty} \sup_{t \leqslant t' \leqslant \lambda t} \frac{s(t') - s(t)}{t^{q-p}} = W_1(\lambda), \tag{9}$$

$$\limsup_{t\to\infty} \sup_{t \leqslant t' \leqslant \lambda t} \frac{s(t') - s(t)}{t'^{q-p}} = W_1^*(\lambda). \tag{9*}$$

Let $s_p(x)$ *be defined for a positive integer* p *as in* (1), *and let*

$$\liminf_{x\to\infty} \frac{s_p(x)}{x^q} = \underline{\sigma}_{p,q}, \quad \limsup_{x\to\infty} \frac{s_p(x)}{x^q} = \bar{\sigma}_{p,q}. \tag{10}$$

Then

$$-\left(\frac{\lambda-1}{p}\right)^p \liminf_{x\to\infty} \frac{s(x)}{x^{q-p}} \leqslant \mathfrak{A}_q(\lambda, p)\,\underline{\sigma}_{p,q} + \mathfrak{B}_q(\lambda, p)\,\bar{\sigma}_{p,q}$$
$$+ \left(\frac{\lambda-1}{p}\right)^{p-1} \int_{1+(1-p^{-1})(\lambda-1)}^{\lambda} W_1(t)dt, \tag{11}$$

where

$$\mathfrak{A}_q(\lambda, p) + \mathfrak{B}_q(\lambda, p) = -\sum_{\nu=0}^{p} (-1)^\nu \binom{p}{\nu} \left\{1 + (p-\nu)\frac{\lambda-1}{p}\right\}^q, \tag{12}$$

$\mathfrak{A}_q(\lambda,p)$ *is the part of the above sum consisting of the negative terms only and* $\mathfrak{B}_q(\lambda,p)$ *is the part of the same sum consisting of the positive terms only.*

Further, for $0 < \theta < 1$, *we have*

$$\left(\frac{1-\theta}{p}\right)^p \limsup_{x\to\infty} \frac{s(x)}{x^{q-p}} \leqslant \mathfrak{C}_q(\theta,p)\,\sigma_{p,q} + \mathfrak{D}_q(\theta,p)\,\bar{\sigma}_{p,q} + \left(\frac{1-\theta}{p}\right)^{p-1} \int_0^{1-(1-p^{-1})(1-\theta)} W_1^*(t)dt, \tag{13}$$

where

$$\mathfrak{C}_q(\theta,p) + \mathfrak{D}_q(\theta,p) = \sum_{\nu=0}^{p} (-1)^\nu \binom{p}{\nu} \left\{1 - \nu\frac{1-\theta}{p}\right\}^q, \tag{14}$$

$\mathfrak{C}_q(\theta,p)$ *is the part of the above sum containing the negative terms alone and* $\mathfrak{D}_q(\theta,p)$ *is the part of the same sum containing the positive terms alone.*

(A condition such as (9) is to be read: "The left-hand member exists as a finite number and equals $W_1(\lambda)$."

(9*) follows from (9) since

$$\frac{s(t') - s(t)}{t'^{q-p}} = \left(\frac{t}{t'}\right)^{q-p} \frac{s(t') - s(t)}{t^{q-p}}$$

Similarly (9) follows from (9*).)

Proof. From Lemma 1 we have at once

$$-h^p \frac{s(x)}{x^q} = -\frac{\Delta_h^p s_p(x)}{x^q} + \int_x^{x+h} dt_1 \int_{t_1}^{t_1+h} dt_2 \cdots \int_{t_{p-1}}^{t_{p-1}+h} \frac{s(t) - s(x)}{x^q} dt.$$

Denoting by I and J the first and the second terms respectively on the right, we can write the above relation as

$$-\left(\frac{h}{x}\right)^p \frac{s(x)}{x^{q-p}} = I + J. \tag{15}$$

In J, t is such that $x \leqslant t_1 \leqslant t \leqslant t_1 + (p-1)h$, and so

$$J \leqslant \int_x^{x+h} \sup_{x\leqslant t\leqslant t_1+(p-1)h} \left\{\frac{s(t) - s(x)}{x^q}\right\} h^{p-1} dt_1.$$

If $h = (\lambda - 1)x/p$ and $xt' = t_1 + (p-1)h$, this gives us

$$J \leqslant \int_{1+(1-p^{-1})(\lambda-1)}^{\lambda} \sup_{x\leqslant t\leqslant xt'} \left\{\frac{s(t) - s(x)}{x^{q-p}}\right\} \left(\frac{h}{x}\right)^{p-1} dt',$$

or, on account of (9),

$$J \leqslant \left(\frac{h}{x}\right)^{p-1} \cdot \int_{1+(1-p^{-1})(\lambda-1)}^{\lambda} W_1(t')dt' + \left(\frac{h}{x}\right)^{p-1} \cdot o(1) \qquad (x \to \infty). \tag{16}$$

Next

$$I = -\sum_{\nu=0}^{p} (-1)^\nu \binom{p}{\nu} \frac{s_p(x + \overline{p-\nu}h)}{(x + \overline{p-\nu}h)^q} \left(1 + \frac{\overline{p-\nu}}{x}h\right)^q \tag{17}$$

where the factor multiplying $s_p(x + \overline{p - \nu}h)/(x + \overline{p - \nu}h)^q$, $\nu = 0,1,2, \ldots p$, is

$$-(-1)^\nu \binom{p}{\nu} \left\{1 + (p - \nu)\frac{\lambda - 1}{p}\right\}^q$$

which is independent of x. Consequently we get, letting $x \to \infty$ in (17) and recalling (10),

$$\limsup_{x\to\infty} I \leqslant \mathfrak{A}_q(\lambda, p)\, \underline{\sigma}_{p,q} + \mathfrak{B}_q(\lambda, p)\, \bar{\sigma}_{p,q} \tag{18}$$

by the definitions of $\mathfrak{A}_q$ and $\mathfrak{B}_q$ which follow (12). Taking upper limits of both sides of (15) as $x \to \infty$ and using (16) and (18), we establish the first conclusion (11).

To prove the second conclusion (13), we get from Lemma 2 the relation

$$k\frac{s(x)}{x^q} = \frac{\Delta_k^p s_p(x)}{x^q} + \int_{x-k}^{x} dt_1 \int_{t_1-k}^{t_1} dt_2 \ldots \int_{t_{p-1}-k}^{t_{p-1}} \frac{s(x) - s(t)}{x^q} dt,$$

and rewrite it, denoting the first and the second terms on its right side by I^* and J^* respectively:

$$\left(\frac{k}{x}\right)^p \frac{s(x)}{x^{q-p}} = I^* + J^*. \tag{19}$$

In J^*, t lies in the interval $t_1 - (p-1)k \leqslant t \leqslant t_1 \leqslant x$, so that

$$J^* \leqslant \int_{x-k}^{x} \sup_{t_1-(p-1)k\leqslant t\leqslant x} \left\{\frac{s(x) - s(t)}{x^q}\right\} k^{p-1} dt_1.$$

If $k = (1 - \theta)x/p$ and $xt' = t_1 - (p-1)k$, we can write the above inequality successively in the forms

$$J^* \leqslant \int_0^{1-(1-p^{-1})(1-\theta)} \sup_{xt'\leqslant t\leqslant x} \left\{\frac{s(x) - s(t)}{x^{q-p}}\right\} \left(\frac{k}{x}\right)^{p-1} dt',$$

$$J^* \leqslant \left(\frac{k}{x}\right)^{p-1} \int_0^{1-(1-p^{-1})(1-\theta)} W_1^*\left(\frac{1}{t'}\right) dt' + \left(\frac{k}{x}\right)^{p-1} \cdot o(1) \qquad (x \to \infty), \tag{20}$$

using (9*). Next

$$I^* = \sum_{\nu=0}^{p} (-1)^\nu \binom{p}{\nu} \frac{s_p(x - \nu k)}{(x - \nu k)^q} \left(1 - \nu\frac{k}{x}\right)^q \tag{21}$$

where the factor multiplying $s_p(x - \nu k)/(x - \nu k)^q$, $\nu = 0,1,2, \ldots p$, is

$$(-1)^\nu \binom{p}{\nu} \left\{1 - \nu\frac{1 - \theta}{p}\right\}^q$$

which is free from x. Therefore we obtain, letting $x \to \infty$ in (21),

$$\limsup_{x\to\infty} I^* \leqslant \mathfrak{C}_q(\theta, p)\, \underline{\sigma}_{p,q} + \mathfrak{D}_q(\theta, p)\, \bar{\sigma}_{p,q} \tag{22}$$

on account of (10) and our definitions of $\mathfrak{C}_q$, $\mathfrak{D}_q$ following (14). By taking

upper limits of both sides of (19) as $x \to \infty$, and using (20) and (22), we immediately get (13).

4. Deductions. The deductions from Theorem I which have been outlined in an earlier section are effected by means of the two simple observations noted below as lemmas.

LEMMA 3. *If* $W_1(\lambda)$ *defined by* (9) *is such that*

$$\lim_{\lambda \to 1+0} W_1(\lambda) \leqslant 0,$$

then $W_1^*(\lambda)$ *defined by* (9*) *is also such that*

$$\lim_{\lambda \to 1+0} W_1^*(\lambda) \leqslant 0,$$

and conversely; further, the integrals in (11) *and* (13) *satisfy the conditions:*

$$\limsup_{\lambda \to 1+0} \frac{p}{\lambda - 1} \int_{1+(1-p^{-1})(\lambda-1)}^{\lambda} W_1(t)dt \leqslant 0,$$

$$\limsup_{\theta \to 1-0} \frac{p}{1-\theta} \int_{\theta}^{1-(1-p^{-1})(1-\theta)} W_1^*(t)dt \leqslant 0.$$

The proof is obvious.

LEMMA 4. *The function of* λ *defined in* (12) *and that of* θ *defined in* (14) *satisfy the conditions:*

$$(23)\quad \mathfrak{A}_q(\lambda,p) + \mathfrak{B}_q(\lambda,p) \sim -\left(\frac{\lambda-1}{p}\right)^p q(q-1)\dots(q-p+1) \quad \text{as } \lambda \to 1+0,$$

$$(24)\quad \mathfrak{C}_q(\theta,p) + \mathfrak{D}_q(\theta,p) \sim \left(\frac{1-\theta}{p}\right)^p q(q-1)\dots(q-p+1) \quad \text{as } \theta \to 1-0.$$

Proof. The proof of (23) is given below; that of (24) is similar.

By (12),

$$\begin{aligned}
\mathfrak{A}_q(\lambda,p) + \mathfrak{B}_q(\lambda,p) &= -\sum_{\nu=0}^{p} (-1)^\nu \binom{p}{\nu} \left\{1 + (p-\nu)\frac{h}{x}\right\}^q \qquad \left(\frac{h}{x} = \frac{\lambda-1}{p}\right) \\
&= -x^{-q}\,\Delta_h^p\, x^q \\
&= -x^{-q}\,h^p \left(\frac{d^p x^q}{dx^p}\right)_{x=\xi} \qquad (x < \xi < x + ph = \lambda x) \\
&= -x^{p-q}\left(\frac{h}{x}\right)^p q(q-1)\dots(q-p+1)\xi^{q-p} \\
&\sim -\left(\frac{\lambda-1}{p}\right)^p q(q-1)\dots(q-p+1) \qquad (\lambda \to 1+0)
\end{aligned}$$

It is clear that, *in the particular case* $q = p$, (23) *and* (24) *reduce to*

$$(23')\qquad \mathfrak{A}_p(\lambda,p) + \mathfrak{B}_p(\lambda,p) = -\left(\frac{\lambda-1}{p}\right)^p p!\,,$$

$$(24')\qquad \mathfrak{C}_p(\theta,p) + \mathfrak{D}_p(\theta,p) = \left(\frac{1-\theta}{p}\right)^p p!\,.$$

To explain the derivation of Theorem A′ from Theorem I, we have only to rewrite the former as follows in the notation of the latter.

COROLLARY I(1). *If, in Theorem* I,

$$\lim_{\lambda\to 1+0} W_1(\lambda) \leqslant 0 \quad \textit{and hence} \quad \lim_{\lambda\to 1+0} W_1^*(\lambda) \leqslant 0,$$

and also

$$\underline{\sigma}_{p,q} = \bar{\sigma}_{p,q} = l,$$

then

$$\lim_{x\to\infty} \frac{s(x)}{x^{q-p}} = q(q-1)\dots(q-p+1)l.$$

For, dividing (11) and (13) throughout by $(\lambda-1)^p/p^p$ and $(1-\theta)^p/p^p$ respectively, and then letting $\lambda \to 1+1-0$, $\theta \to 0$, we get as a result of Lemmas 3 and 4,

$$-\liminf_{x\to\infty} \frac{s(x)}{x^{q-p}} \leqslant -q(q-1)\dots(q-p+1)l,$$

$$\limsup_{x\to\infty} \frac{s(x)}{x^{q-p}} \leqslant q(q-1)\dots(q-p+1)l,$$

which together imply the conclusion of Corollary I(1).

If $q = p$ in Theorem I, we find from (9), (9*) and (8) that

$$W_1(\lambda) = W_1^*(\lambda) = w_1(\lambda),$$

and from (3) and (10) that

$$\underline{\sigma}_{p,p} = \underline{C}_p/p! \quad , \quad \bar{\sigma}_{p,p} = \bar{C}_p/p! \quad .$$

Hence, when $q = p$ in Theorem I, the result is the following extension of Theorem D obtained by me some time ago (**9**, Theorem A).

COROLLARY I(2). *If $s(x)$ is integrable in every finite interval of $x > 0$ and such that, for $\lambda > 1$,*

$$\limsup_{t\to\infty} \sup_{t\leqslant t'\leqslant t} \{s(t') - s(t)\} = w_1(\lambda), \tag{8}$$

then, for $0 < \theta < 1 < \lambda$,

$$-\left(\frac{\lambda-1}{p}\right)^p \underline{C}_0 \leqslant \frac{\mathfrak{A}_p(\lambda,p)\underline{C}_p + \mathfrak{B}_p(\lambda,p)\bar{C}_p}{p!} + \left(\frac{\lambda-1}{p}\right)^{p-1} \int_{1+(1-p^{-1})(\lambda-1)}^{\lambda} w_1(t)dt, \tag{11′}$$

$$\left(\frac{1-\theta}{p}\right)^p \bar{C}_0 \leqslant \frac{\mathfrak{C}_p(\theta,p)\underline{C}_p + \mathfrak{D}_p(\theta,p)\bar{C}_p}{p!} + \left(\frac{1-\theta}{p}\right)^{p-1} \int_{0}^{1-(1-p^{-1})(1-\theta)} w_1(t)dt, \tag{13′}$$

where $\mathfrak{A}_p$, $\mathfrak{B}_p$, $\mathfrak{C}_p$, $\mathfrak{D}_p$ are obtained with $q = p$ in $\mathfrak{A}_q$, $\mathfrak{B}_q$, $\mathfrak{C}_q$, $\mathfrak{D}_q$ respectively as defined immediately after (12) *and* (14).

In the particular case in which the hypothesis is

$$\lim_{\lambda \to 1+0} w_1(\lambda) < 0 \quad , \quad \underline{C}_p = \bar{C}_p = l \quad ,$$

inequalities (11′) *and* (13′) *together reduce to the conclusion:*

$$\underline{C}_0 = \bar{C}_0 = l, \quad i.e. \quad \lim_{x \to \infty} s(x) = l,$$

on account of (23′), (24′) and Lemma 3 with $q = p$.

Theorem B is thus a particular case of Corollary I(2). Theorem C is a similar particular case of the next corollary got by making a small change in the proof of (11′) of Corollary I(2).

COROLLARY I(3). *If* $\lambda > 1$ *and*

$$\limsup_{t \to \infty} \sup_{t \leqslant t' \leqslant \lambda t} \left| \frac{1}{(\lambda - 1)t} \int_t^{t'} \{s(u) - s(t)\} du \right| = \Omega(\lambda), \tag{25}$$

then

$$-\left(\frac{\lambda - 1}{p}\right)^p \underline{C}_0 < \frac{\mathfrak{A}_p(\lambda, p)\underline{C}_p + \mathfrak{B}_p(\lambda, p)\bar{C}_p}{p!} + 2\left(\frac{\lambda - 1}{p}\right)^p p\Omega(\lambda), \tag{26}$$

and there is a similar inequality with $\bar{C}_0$, $-\bar{C}_p$, $-\underline{C}_p$ *taking the places of* $-\underline{C}_0$, $\underline{C}_p$, $\bar{C}_p$ *respectively.*

In the particular case in which the hypothesis is that of Theorem C, *that is,*

$$\lim_{\lambda \to 1+0} \Omega(\lambda) = 0 \quad , \quad \underline{C}_p = \bar{C}_p = l,$$

inequality (26) *and its companion specified after it together yield the conclusion of Theorem* C, *viz.*

$$\underline{C}_0 = \bar{C}_0 = l,$$

as a result of (23′).

To prove Corollary I(3) in all its generality, we write down (15) with $q = p$ and find an upper estimate for J, using the following consequence of (25):

$$\int_{t_{p-1}}^{t_{p-1}+h} \{s(t) - s(x)\} dt < \left| \int_x^{t_{p-1}} \{s(t) - s(x)\} dt \right| + \left| \int_x^{t_{p-1}+h} \{s(t) - s(x)\} dt \right|$$
$$< 2\{\Omega(\lambda) + o(1)\}(\lambda - 1)x \qquad (x \to \infty).$$

From this we obtain in succession

$$J < x^{-p} \int_x^{x+h} dt_1 \int_{t_1}^{t_1+h} dt_2 \ldots \int_{t_{p-2}}^{t_{p-2}+h} 2\{\Omega(\lambda) + o(1)\}(\lambda - 1)x dt_1 \qquad (x \to \infty),$$

$$\limsup_{x \to \infty} J < 2\left(\frac{\lambda - 1}{p}\right)^{p-1} (\lambda - 1)\Omega(\lambda),$$

finally reaching (26) by a repetition of the rest of the argument used to prove (11). (26) has a companion as stated, resulting from the replacement of $s(x)$

by $-s(x)$ which is obviously permissible in our hypothesis (25) and all arguments therefrom.

In the case $p = 1$, Corollary I(3) can be modified to become a slight extension and simplification of Pitt's theorem already referred to (7, Theorem 1). This modification of Corollary I(3), analogous to Theorem D, is stated below.

COROLLARY I(4). *If, given some $\lambda > 1$, we can find, corresponding to every sufficiently large t, $R = R(t)$ tending to λ as $t \to \infty$ and such that*

$$\limsup_{t\to\infty} \left| \frac{1}{(R-1)t} \int_t^{Rt} \{s(u) - s(t)\}du \right| = \omega(\lambda), \tag{27}$$

then

$$-(\lambda - 1)\underline{C}_0 \leqslant -\lambda \underline{C}_1 + \bar{C}_1 + (\lambda - 1)\omega(\lambda), \tag{28}$$

$$(\lambda - 1)\bar{C}_0 \leqslant -\underline{C}_1 + \lambda \bar{C}_1 + (\lambda - 1)\omega(\lambda). \tag{29}$$

In particular, when (27) is simply

$$\lim_{t\to\infty} \left| \frac{1}{(R-1)t} \int_t^{Rt} \{s(u) - s(t)\}du \right| = 0, \tag{27'}$$

and $\underline{C}_1 = \bar{C}_1$, (28) and (29) together reduce to the equality

$$\underline{C}_0 = \bar{C}_0.$$

Corollary I(4) is proved from the following relation which is the case $q = p = 1$ of (15) with $h = (R-1)x$ and $R = R(x)$:

$$-(R-1)s(x) = -RC_1(Rx) + C_1(x) + \frac{1}{x}\int_x^{Rx} \{s(t) - s(x)\}dt.$$

Taking upper limits of both sides as $x \to \infty$ and using (27), we get at once (28) and deduce (29) from it by changing $s(x)$ to $-s(x)$, such a change being permissible in (27) and arguments based thereon.

REMARK ON CONDITION (27'). *This Tauberian condition, like Pitt's more complicated form of it (7), though sufficient to make the convergence of $s(x)$ follow from the (C, 1) summability of $s(x)$, is not always sufficient to make the convergence of $s(x)$ follow from the Abel summability of $s(x)$.*

(Pitt has, instead of (27'), the more complicated condition: *given $\epsilon > 0$, we can find $\eta(\epsilon) > 0$, $R = R(t, \epsilon)$ corresponding to every sufficiently large t, so that*

$$R > 1 + \eta, \quad \left| \int_T^{RT} \{s(u) - s(t)\}du \right| < (R-1)T\epsilon$$

for some $T = T(\epsilon, t)$ satisfying $tR^{-1} < T < t$.)

Pitt's example itself (7, Theorem 2) serves to establish this fact. The example is of a non-convergent $s(x)$ which is Abel summable and defined as follows:

$s(x) = (-2)^m$ for $\lambda_m \leqq x < \lambda_{m+1}$, $\lambda_m = (2m+1)\log(2m+1)$, $m = 0, 1, 2, \ldots$. Pitt's discussion shows that, for this $s(x)$ there is an $R = R(t)$ corresponding to every sufficiently large t, such that $R(t) \to \lambda$ as $t \to \infty$ and (27′) is fulfilled in the form

$$\frac{1}{(R-1)t}\int_t^{Rt} \{s(u) - s(t)\}du = 0.$$

(What Pitt has actually proved is that, corresponding to every sufficiently large t, $\lambda_M \leqq t < \lambda_{M+1}$, we can find $R = R(M)$ tending to $\lambda = 2$ as $t \to \infty$, so that

$$\frac{1}{(R-1)\lambda_M}\int_{\lambda_M}^{R\lambda_M} \{s(u) - s(\lambda_M)\}du = 0.$$

However, it is easy to show that Pitt's result remains true when $\lambda = 2$ is replaced by any $\lambda > 1$, λ_M by t and $R = R(M)$ by another $R = R(t)$.)

5. A supplementary result. To make this study complete, a complement to Theorem I under a two-sided Tauberian condition is proved below. This complement, in the special case $q = p$, reduces to a result previously obtained by me (**9**, Theorem B), and, in the further special case $q = p = 1$, to Karamata's complement to Theorem D (**3**, Satz 1, second part) under the condition (8) together with a similar condition on $-s(x)$ instead of $s(x)$.

THEOREM II. *If, in Theorem* I, *we are given, in addition to either* (9) *or* (9*), *one of the following conditions which necessarily involves the other:*

$$\liminf_{t\to\infty} \inf_{t\leqq t'\leqq \lambda t} \frac{s(t') - s(t)}{t^{q-p}} = -W_2(\lambda), \tag{30}$$

$$\liminf_{t\to\infty} \inf_{t\leqq t'\leqq \lambda t} \frac{s(t') - s(t)}{t'^{q-p}} = -W_2^*(\lambda), \tag{30*}$$

we shall have, in addition to (11) *and* (13),

$$-\left\{\left(\frac{\lambda-1}{p}\right)^p + \left(\frac{1-\theta}{p}\right)^p\right\} \liminf_{x\to\infty} \frac{s(x)}{x^{q-p}}$$

$$\leqq \begin{cases} \left\{\mathfrak{A}_q(\lambda,p) - \mathfrak{D}_q(\theta,p)\right\}\underline{\sigma}_{p,q} + \left\{\mathfrak{B}_q(\lambda,p) - \mathfrak{C}_q(\theta,p)\right\}\bar{\sigma}_{p,q} \\ \qquad + \left\{\dfrac{1+(-1)^{p-1}}{2}\right\}(\underline{\sigma}_{p,q} - \bar{\sigma}_{p,q}) \\ +\left(\dfrac{\lambda-1}{p}\right)^{p-1}\displaystyle\int_{1+(1-p^{-1})(\lambda-1)}^{\lambda} W_1(t)dt + \left(\dfrac{1-\theta}{p}\right)^{p-1}\int_0^{1-(1-p^{-1})(1-\theta)} W_2^*(1/t)dt \end{cases} \tag{31}$$

and a similar inequality with lim sup $s(x)/x^{q-p}$ *in place of* $-$ lim inf $s(x)/x^{q-p}$ *deduced from* (31) *by taking* $-s(x)$ *in place of* $s(x)$.

Proof. By combining (15) and (19), we get

$$-\left(\frac{h}{x}\right)^p \frac{s(x)}{x^{q-p}} - \left(\frac{k}{x}\right)^p \frac{s(x)}{x^{q-p}} = I + J - I^* - J^*. \tag{32}$$

Taking $h = (\lambda - 1)x/p$, $k = (1 - \theta)x/p$, and arguing as in the derivation of (16) and (20), we obtain

$$\begin{aligned}\limsup_{x\to\infty} (J - J^*) &\leqslant \limsup_{x\to\infty} J + \limsup_{x\to\infty} (-J^*) \\ &\leqslant \left(\frac{\lambda - 1}{p}\right)^{p-1} \int_{1+(1-p^{-1})(\lambda-1)}^{\lambda} W_1(t)dt + \\ &\qquad \left(\frac{1-\theta}{p}\right)^{p-1} \int_{\theta}^{1-(1-p^{-1})(1-\theta)} W_2^*(1/t)dt.\end{aligned} \tag{33}$$

We have also, from the expression for I in (17) and that for I^* in (21),

$$\limsup_{x\to\infty} (I - I^*) \leqslant \begin{cases} \mathfrak{A}_q \sigma_{p,q} + \mathfrak{B}_q \bar{\sigma}_{p,q} - \mathfrak{C}_q \bar{\sigma}_{p,q} - \mathfrak{D}_q \sigma_{p,q} & \text{if } p \text{ is even,} \\ \mathfrak{A}_q \sigma_{p,q} + (\mathfrak{B}_q - 1)\bar{\sigma}_{p,q} - \mathfrak{C}_q \bar{\sigma}_{p,q} - (\mathfrak{D}_q - 1)\sigma_{p,q} & \text{if } p \text{ is odd,} \end{cases} \tag{34}$$

where the distinction between the cases of odd p and even p arises thus. If p is odd and only then, the last term in I is $s_p(x)/x^q$ and this cancels out the first term in $-I^*$ which is in any case $-s_p(x)/x^q$; the result is that the contribution (arising from I) to the positive terms which make up $\mathfrak{B}_q$ is less than what the form of I suggests, by 1, and the contribution (arising from $-I^*$) to the negative terms which make up $-\mathfrak{D}_q$ is more than what the form of $-I^*$ suggests, by 1. (31) follows from (32), (33) and (34).

6. Concluding remarks. There is a special case of interest in the results of this note, when $s(x)$ is, as in Pitt's example, a λ_n − step function with steps at points of any sequence $\{\lambda_n\}$ such that

$$0 < \lambda_0 < \lambda_1 < \ldots, \lambda_n \to \infty,$$

that is,

$$s(x) = \begin{cases} a_0 + a_1 + \ldots + a_n & \text{for } \lambda_n \leqslant x < \lambda_{n+1}, \quad n \geqslant 0, \\ 0 & \text{for } 0 \leqslant x < \lambda_0. \end{cases}$$

In this case, the (C, α) summability of $s(x)$ becomes the summability of Σa_n by Riesz means of order α and type (λ_n), usually called (R, λ_n, α) summability; and Corollary I(2) can be used, as elsewhere **(10; 11)**, to extend certain Tauberian theorems of G. Ricci's for Σa_n summable to l by the method of Dirichlet's series or the (A, λ_n) method, that is, Σa_n such that

$$\sum_{n=0}^{\infty} a_n e^{-\lambda_n s} \text{ converges for } s > 0 \text{ and tends to } l \text{ as } s \to +0.$$

An open question **(10, §1.1)** which may be recalled in this context is *whether the following theorem for* (R, λ_n, α) *summability is one possessing no precise analogue for* (A, λ_n) *summability*, i.e. one belonging possibly to a class of Tauberian theorems peculiar to Cesàro summability like the particular case of Corollary I(4).

THEOREM X. *If*

(i) $$\sum_{n=0}^{\infty} a_n$$

is (R, λ_n, α) *summable to* l *for some* $\alpha > 0$,

(ii) $$\lim_{\lambda \to 1+0} \limsup_{n \to \infty} \max_{\lambda_n < \lambda_m < \lambda\lambda_n} (a_{n+1} + a_{n+2} + \ldots + a_m) \leqslant 0,$$

then

$$\liminf_{n \to \infty} (a_0 + a_1 + \ldots + a_n) = l.$$

Theorem X (**10**, Theorem f) is a simple consequence of Corollary I(2), and it has the imperfect analogue for (A, λ_n) summability, stated below, whose special case $\alpha = 0$ follows from a reformulation of one of Ricci's theorems (**10**, Theorem G) and every case $\alpha > 0$ follows from Theorem X and my generalization (**10**, Lemma 2) of a theorem due to O. Szász.

THEOREM Y. *Theorem* X *can be restated with* (i) *replaced by the* (A, λ_n) *summability of* Σa_n *to* l *and* (ii) *augmented by the condition that, for some* $\alpha > 0$,

$$\sum_{\lambda_\nu < x} (x - \lambda_\nu)^\alpha a_\nu \lambda_\nu = O_R(x^{\alpha+1}) \qquad (x \to \infty).$$

REFERENCES

1. L. S. Bosanquet, *Note on convexity theorems*, J. London Math. Soc., *18* (1943), 239–248.
2. G. Doetsch, *Ueber die Cesàrosche Summabilität bei Reihen und eine Erweiterung des Grenswerthegriffs bei integrablen Funktionen*, Math. Z., *11* (1921), 161–179.
3. J. Karamata, *Beziehungen Zwischen den Oscillationsgrenzen einer Funktion und ihrer arithmetischen Mittel*, Proc. London Math. Soc., (2) *43* (1937), 20–25.
4. ———, *Quelques théorèmes inverses relatifs aux procédés de sommabilitié de Cesàro et Riesz*, Acad. Serbe Sci. Publ. Inst. Math., *3* (1950), 53–71.
5. N. Obrechkoff, *Sur un formule pour les différences divisées et sur les limites de fonctions et de leurs dérivées*, C.R. Acad. Bulgare Sci., *2* (1949), 5–8.
6. M. Parthasarathy and C. T. Rajagopal, *A theorem on the Riemann-Liouville integral*, Math. Z., *55* (1951), 84–91.
7. H. R. Pitt, *A note on Tauberian conditions for Abel and Cesàro summability*, Proc. Amer. Math. Soc., *6* (1955), 616–619.
8. C. T. Rajagopal, *A note on the oscillation of Riesz means of any order*, J. London Math. Soc., *21* (1946), 275–282.
9. ———, *On the limits of oscillation of a function and its Cesàro means*, Proc. Edinburgh Math. Soc., (2) *7* (1946), 162–167.
10. ———, *On some extensions of Ananda Rau's converse of Abel's theorem*, J. London Math. Soc., *23* (1948), 38–44,
11. ———, *On a Tauberian theorem of G. Ricci*, Proc. Edinburgh Math. Soc., (2) *8* (1949), 143–146.
12. ———, *On Tauberian theorems for the Riemann-Liouville integral*, Acad. Serbe Sci. Publ. Inst. Math., *6* (1954), 27–46.
13. D. V. Widder, *The Laplace Transform* (Princeton, 1941).

Ramanujan Institute of Mathematics
Madras, India

ON THE COMPLEMENTARY FUNCTIONS OF THE FRESNEL INTEGRALS

ERWIN KREYSZIG

1. Introduction. As is well known, the functions

$$c(u) = \int_u^\infty \cos(p^2)dp, \quad s(u) = \int_u^\infty \sin(p^2)dp \tag{1.1}$$

have various applications in theoretical physics and engineering. It is thus worthwhile to study their behaviour for real and complex values of the argument. Since

$$c(u) = \tfrac{1}{2}\int_{u^2}^\infty t^{-\frac{1}{2}}\cos t\,dt, \quad s(u) = \tfrac{1}{2}\int_{u^2}^\infty t^{-\frac{1}{2}}\sin t\,dt$$

we may consider the functions

$$c(z) = \int_z^\infty t^{-\frac{1}{2}}\cos t\,dt, \qquad s(z) = \int_z^\infty t^{-\frac{1}{2}}\sin t\,dt, \quad z = x + iy, \tag{1.2}$$

instead of (1.1).

We have

$$c(z) = C - C(z), \quad s(z) = S - S(z) \tag{1.3}$$

where

$$C(z) = \int_0^z t^{-\frac{1}{2}}\cos t\,dt, \qquad S(z) = \int_0^z t^{-\frac{1}{2}}\sin t\,dt \tag{1.4}$$

are the Fresnel integrals and

$$C = \lim_{z\to\infty} C(z) = \sqrt{\frac{\pi}{2}}, \qquad S = \lim_{z\to\infty} S(z) = \sqrt{\frac{\pi}{2}}. \tag{1.5}$$

In order that the relation (1.3) be valid, one has to choose a path of integration which goes asymptotically parallel to the x-axis to infinity. By means of (1.3) results about $c(z)$ and $s(z)$ can be obtained from recent considerations (1) of the Fresnel integrals; since these consequences are immediate we shall not consider them in detail.

2. Relations to other known functions. The functions $c(z)$ and $s(z)$ can be represented by certain $W_{k,m}$-functions. We have

$$W_{-1/4,1/4}(\pm iz) = e^{\mp iz/2}(\pm iz)^{-1/4}\int_0^\infty \left(1 \pm \frac{t}{iz}\right)^{-\frac{1}{2}} e^{-t}dt. \tag{2.1}$$

Received September 15, 1956.

Setting $1 + t/iz = w/z$ and $1 - t/iz = w/z$, respectively, we obtain

$$W_{-1/4,1/4}(\pm iz) = z^{1/4}e^{\pm i(z/2+3\pi/8)}\int_z^\infty w^{-\frac{1}{2}}e^{\mp tw}\,dw \qquad \left(\arg z \neq \frac{\pi}{2}\right),$$

where one has to integrate along the real axis. Hence we find

$$\begin{aligned} c(z) &= 2^{-1}z^{-1/4}\{e^{-i\alpha(z)}W_{-1/4,1/4}(iz) + e^{i\alpha(z)}W_{-1/4,1/4}(-iz)\}, \\ s(z) &= 2^{-1}z^{-1/4}\{e^{-i\beta(z)}W_{-1/4,1/4}(iz) + e^{i\beta(z)}W_{-1/4,1/4}(-iz)\}, \end{aligned} \qquad \left(\arg z \neq \frac{\pi}{2}\right) \tag{2.2}$$

where

$$\alpha(z) = \tfrac{1}{2}\left(z + \frac{3\pi}{4}\right), \qquad \beta(z) = \tfrac{1}{2}\left(z - \frac{\pi}{4}\right).$$

For both integrals on the right hand side of (2.2) one has to choose a common path of integration which coincides asymptotically with the real axis.

Using the error function

$$\mathrm{Erf}(z) = \int_z^\infty e^{-t^2}dt$$

we find

$$\begin{aligned} c(z) &= i^{\frac{1}{2}}\mathrm{Erf}(\sqrt{-iz}) + i^{-\frac{1}{2}}\mathrm{Erf}(\sqrt{iz}), \\ s(z) &= i^{-\frac{1}{2}}\mathrm{Erf}(\sqrt{-iz}) + i^{\frac{1}{2}}\mathrm{Erf}(\sqrt{iz}), \end{aligned} \qquad \left(\arg z \neq \frac{\pi}{2}\right). \tag{2.3}$$

Relations between $c(z)$, $s(z)$ and the incomplete Gamma function $Q(z, \beta)$ have been mentioned in **(1)**. There exist also relations to Lommel's functions which can be obtained from the representations given in **(4)**.

3. Improvement of the accuracy of the asymptotic expansion. From **(1)** we obtain the following asymptotic expansions valid for all complex arguments with the exception of purely imaginary ones:

$$\begin{aligned} c(z) &\sim -z^{-\frac{1}{2}}\,(a(z)\cos z + b(z)\sin z) \\ s(z) &\sim \quad z^{-\frac{1}{2}}\,(b(z)\cos z - a(z)\sin z) \end{aligned} \tag{3.1}$$

where

$$a(z) = \sum_{m=1}^{\infty}(-1)^m\frac{1.3\ldots(4m-3)}{(2z)^{2m-1}}, \quad b(z) = 1 + \sum_{m=1}^{\infty}(-1)^m\frac{1.3\ldots(4m-1)}{(2z)^{2m}}.$$

For real values $z = x$ the optimal accuracy of (3.1) can be considerably improved by summing the divergent part of (3.1) by Euler's method; we have to multiply the smallest term of the series by the function obtained through that Euler summation process. In detail: The term of $a(z)$ which has the smallest absolute value corresponds to the largest value of m for which

$$4m^2 - \tfrac{1}{4} < |z|^2. \tag{3.2}$$

Let us denote this term by $a_{m_0}(z)$. We first consider values of x which are integers. Equation (3.2) is valid for $x = 2m$. Then we have

$$a(x) = a_1(x) + \ldots + a_{m_0}(x)K_I(x)$$

where

$$(3.3) \qquad K_I(x) = 1 - \frac{(2x-1)(2x+1)}{(2x)^2} + \frac{(2x-1)(2x+1)\ldots(2x+5)}{(2x)^4} - + \ldots$$

or

$$K_I(x) = 1 - \left(1 - \frac{1}{(2x)^2}\right) + \left(1 + \frac{8}{2x} + \frac{14}{(2x)^2} - \frac{8}{(2x)^3} - \frac{15}{(2x)^4}\right) - + \ldots$$

$$= (1 - 1 + 1 - + \ldots) + (8 - 24 + 48 - + \ldots)\frac{1}{2x}$$

$$+ (1 + 14 - 205 + 924 - + \ldots)\frac{1}{(2x)^2} + \ldots.$$

The sequence of terms occurring in the coefficient of $(2x)^{-p}$, $p = 0, 1, \ldots$, is such that Euler's summation always yields a finite expression for each of these coefficients. In this way we find

$$(3.4) \qquad K_I(x) = 0.5 + \frac{1}{2x} - \frac{1.75}{(2x)^2} + \frac{4.5}{(2x)^3} - \frac{3.875}{(2x)^4} - \frac{146}{(2x)^5} + \ldots.$$

Setting

$$b(x) = b_1(x) + \ldots + b_{m_0-1}(x) + b_{m_0}(x) + \ldots$$

or

$$b(x) = b_1(x) + \ldots + b_{m_0-1}(x) + a_{m_0}(x)K_{II}(x)$$

and $x = 2m$, in consequence of

$$b_m(x) = a_m(x)\frac{4m-1}{2x}, \qquad m = 1,2,\ldots,$$

we obtain

$$K_{II}(x) = \frac{2x-1}{2x} - \frac{(2x-1)(2x+1)(2x+3)}{(2x)^3} + \ldots.$$

If we sum this expression by a method analogous to that described above we find

$$(3.5) \qquad K_{II}(x) = 0.5 - \frac{0.5}{2x} - \frac{0.75}{(2x)^2} + \frac{6.25}{(2x)^3} - \frac{44.375}{(2x)^4} + \ldots.$$

The term of $b(z)$ which has the smallest absolute value corresponds to the largest value of m for which

$$(3.6) \qquad (4m+1)(4m+3) < 4|z|^2.$$

Let us denote this term of $b(z)$ by

$$b_{m_0^*}(z).$$

Again we consider real values $z = x$. Equation (3.6) is valid for $x = 2m + 1$. Then we have

$$b(x) = b_1(x) + \ldots + b_{m_0^*}(x)K_1(x)$$

and, since

$$a_{m+1}(x) = -b_m(x)\frac{4m+1}{2x}, \qquad m = 0,1,\ldots,$$

for that value of x we also have

$$a(x) = a_1(x) + \ldots + a_{m_0^*}(x) + a_{m_0^*+1}(x) + \ldots$$
$$= a_1(x) + \ldots + a_{m_0^*}(x) - b_{m_0^*}K_{II}(x),$$

where $K_I(x)$ and $K_{II}(x)$ are given by (3.4) and (3.5), respectively. For non-integer values of x one has to proceed similarly. Using this method the error in the values of the functions for arguments between 3 and 4, say, is about 2–3 per cent of the smallest error obtained by applying (3.1) in the usual way; the relative improvement increases with increasing argument.

4. Properties of the Zeros. In contrast to the Fresnel integrals, the complementary functions $c(z)$ and $s(z)$ have real zeros but no complex ones.

LEMMA 1. *The function $c(z)$ possesses exactly one zero in each of the intervals $J_n : n\pi \leqslant x \leqslant (2n+1)\pi/2$ $(n = 0, 1, \ldots)$ of the real axis; the other intervals, $K_n : (2n+1)\pi/2 < x < (n+1)\pi$ $(n = 0, 1, \ldots)$, do not contain zeros. The function $s(z)$ has exactly one zero in every interval K_n but no zeros in the intervals J_n.*

Proof. We consider $c(z)$. For large values of x the statement is a consequence of (3.1); a certain zero lies at a distance α_n from $n\pi$,

$$x_n = n\pi + \alpha_n, \qquad \alpha_n > 0,$$

where α_n tends to zero if n tends to infinity. Furthermore, since $\sqrt{x}$ is monotone we obtain from the form of the integrand of $c(z)$ that $\alpha_{n-1} > \alpha_n$ where

$$x_{n-1} = (n-1)\pi + \alpha_{n-1}$$

is the preceding zero of $c(z)$. Moreover, from this property the existence and above indicated position of the smaller zeros of $c(z)$ follows. In consequence of Rolle's theorem we have

$$\alpha_n < \frac{\pi}{2}, \qquad n = 0,1,\ldots,$$

and from these inequalities we conclude the existence of intervals which cannot contain zeros. The statement for $s(z)$ can be proved by a similar argument. All zeros are simple.

Using Lemma 1 we obtain

THEOREM 2. *The functions $c(z)$ and $s(z)$ do not have complex zeros.*

Proof. We consider $c(z)$. Since this function is real for real values of the argument we may restrict ourselves to positive values of y. Furthermore we consider only positive values of x; for negative values of x the proof is similar. From (1.2) we have

$$c(x+iy) = c(x) + L(z) \tag{4.1}$$

where

$$L(z) = -i\int_0^y (x+iw)^{-\frac{1}{2}}\cos(x+iw)dw.$$

Let us denote by x_1 and x_2 $(> x_1)$ the real zeros of $c(z)$ contained in the strip S_n: $2n\pi \leqslant x \leqslant (2n+2)\pi$ of the z-plane. For

$$\tfrac{1}{2}(4n+1)\pi \leqslant x \leqslant (2n+1)\pi, \quad \tfrac{1}{2}(4n+3)\pi \leqslant x \leqslant (2n+2)\pi$$

the imaginary part of $L(z)$ has constant sign; for $x_1 \leqslant x \leqslant \frac{1}{2}(4n+1)\pi$ and $x_2 \leqslant x \leqslant \frac{1}{2}(4n+3)\pi$ the real part of $c(z)$ has constant sign. Consequently, if there were complex zeros of $c(z)$ in S_n their real parts would lie in the intervals

$$2n\pi \leqslant x \leqslant x_1 \quad \text{or} \quad (2n+1)\pi \leqslant x \leqslant x_2. \tag{4.2}$$

Setting $(x+iw)^{-\frac{1}{2}} = a + ib$ we have

$$\Im\, L(z) = -\sin x\int_0^y b\sinh w\,dw - \cos x\int_0^y a\cosh w\,dw.$$

Since $x > 0$ and $y > 0$ we have $-b < a$. Furthermore, since

$$\sin x \leqslant \cos x \qquad (2n\pi \leqslant x \leqslant (2n+\tfrac{1}{4})\pi),$$

for these values of x we have $\Im\, L(z) < 0$. By similar reasoning we find $\Im\, L(z) > 0$ for values of z which satisfy $(2n+1)\pi \leqslant x \leqslant (2n+\frac{5}{4})\pi$. Hence the corresponding strips of S_n cannot contain complex zeros of $c(z)$. We prove finally that these strips contain the strips defined by (4.2). We have to show that

$$\alpha_n \equiv x_1 - 2n\pi < \tfrac{1}{4}\pi \quad , \quad \alpha_{n+1} \equiv x_2 - (2n+1)\pi < \tfrac{1}{4}\pi.$$

Since α_n is a monotone decreasing function of n (cf. the proof of Lemma 1) it suffices to prove that $\alpha_0 < \frac{1}{4}\pi$. From (1.3)–(1.5) we have

$$c(\tfrac{1}{4}\pi) = \sqrt{(\tfrac{1}{2}\pi)} - \int_0^{\frac{1}{4}\pi} t^{-\frac{1}{2}}\cos t\,dt.$$

If $0 \leqslant t \leqslant \frac{1}{4}\pi$ then $\cos t \geqslant 2^{-\frac{1}{2}}$ and consequently

$$c(\tfrac{1}{4}\pi) < \sqrt{(\tfrac{1}{2}\pi)} - 2^{-\frac{1}{2}}\int_0^{\frac{1}{4}\pi} t^{-\frac{1}{2}}\,dt = 0.$$

Since $c(x)$ is continuous, $c(0) = \sqrt{(\frac{1}{2}\pi)} > 0$, and $c(\frac{1}{4}\pi) < 0$, the interval $(0, \frac{1}{4}\pi)$ of the real axis must contain a zero of $c(z)$; hence $\alpha_0 < \frac{1}{4}\pi$. This completes the proof of Theorem 2 for $c(z)$. The statement for $s(z)$ can be proved in a similar way.

5. Computation of the Zeros. As follows from §4 approximate values for the zeros (except for the smallest zero of $c(z)$) can be obtained from (3.1). We have

$$c(x) \sim -x^{-\frac{1}{4}} \sin x + \tfrac{1}{2} x^{-3/2} \cos x + \ldots = 0.$$

Hence the first approximation is given by

$$x_{1,n} = n\pi, \qquad n = 1, 2, \ldots.$$

Setting

$$f(x) = -x^{-\frac{1}{4}} \sin x + \tfrac{1}{2} x^{-3/2} \cos x$$

we obtain

$$f(x_{1,n}) = (-1)^n 2^{-1} (n\pi)^{-3/2}$$

and

$$f'(x_{1,n}) = (-1)^{n+1} (n\pi)^{-\frac{1}{4}} + O((n\pi)^{-5/2}).$$

Hence the second approximation is given by

$$x_{2,n} = n\pi + (2n\pi)^{-1}, \qquad n = 1, 2, \ldots. \tag{5.1}$$

Similarly, we obtain for the zeros of $s(z)$ the second approximation

$$x^*_{2,n} = \tfrac{1}{2}(2n+1)\pi + ((2n+1)\pi)^{-1}, \qquad n = 0, 1, \ldots. \tag{5.2}$$

From (5.1) we obtain the second zero ($n = 1$) of $c(z)$ within an error of 1 per cent and the higher zeros more accurately. From (5.2) the first zero ($n = 0$) of $s(z)$ can be obtained within an error of 7 per cent, the second zero ($n = 1$) within an error of 0.3 per cent, etc.

For a more exact determination of the zeros we can use either the Taylor series development of the Fresnel integrals at $z = 0$, cf. **(1)**, in connection with (1.3) or, if n is large enough, the asymptotic expansion (3.1). In the latter case the method corresponds to that described in **(1, § 7)**. Some of the quotients occurring in the procedure for $c(z)$ involve the function $\tan x_n$. Setting

$$\tan x_n \approx \tan x_{2,n} = \tan (2n\pi)^{-1} \approx (2n\pi)^{-1}$$

we obtain the third approximation for the zeros of $c(z)$ in the form

$$x_{3,n} = n\pi + (2n\pi)^{-1} + 3(2n\pi)^{-3} \tag{5.3}$$

and the higher approximations in the form

$$\begin{aligned} x_{2q,n} &= n\pi + (2n\pi)^{-1} + \sum_{p=2}^{q} (-1)^{p+1} 1.3 \ldots (4p-7)(4p-5)(4p-4)(2n\pi)^{1-2p} \\ x_{2q+1,n} &= x_{2q,n} + (-1)^{q+1} 1.3 \ldots (4q-1)(2n\pi)^{-2q-1}, \qquad q = 2,3, \ldots. \end{aligned} \tag{5.4}$$

Similarly, for the zeros x_n^* of $s(z)$ we obtain the third approximation

$$x^*_{3,n} = \tfrac{1}{2}(2n+1)\pi + ((2n+1)\pi)^{-1} + 3((2n+1)\pi)^{-3} \tag{5.5}$$

and the higher approximations

$$x^*_{2q,n} = \tfrac{1}{2}(2n+1)\pi + ((2n+1)\pi)^{-1} +$$

(5.6)
$$\sum_{p=2}^{q} (-1)^{p+1} 1.3 \ldots (4p-7)\,(4p-5)\,(4p-4)\,((2n+1)\pi)^{1-3p}$$

$$x^*_{2q+1,n} = x_{2q,n} + (-1)^{q+1} 1.3 \ldots (4q-1)\,((2n+1)\pi)^{-2q-1},\ q = 2,3,\ldots.$$

6. Modulus surfaces of $c(z)$ and $s(z)$, tables of zeros. Figure 1 represents the surface $W = |c(z)|$ in three-dimensional xyW-space. Figure 2 shows the surface $W = |s(z)|$ in this space. These graphs yield a clear picture of the distribution of the function values of $c(z)$ and $s(z)$ for complex values of the argument. The surfaces can be investigated by means of differential geometry in a way similar to that used in **(1)** in the study of the Fresnel integrals $C(z)$ and $S(z)$.

The tables contain the first 25 zeros of $c(z)$ and $s(z)$. The calculation was done by means of the methods developed in the preceding sections. Tables of values of the functions for complex values of the argument are obtained from **(2; 3)** by means of the relation (1.3) of this paper.

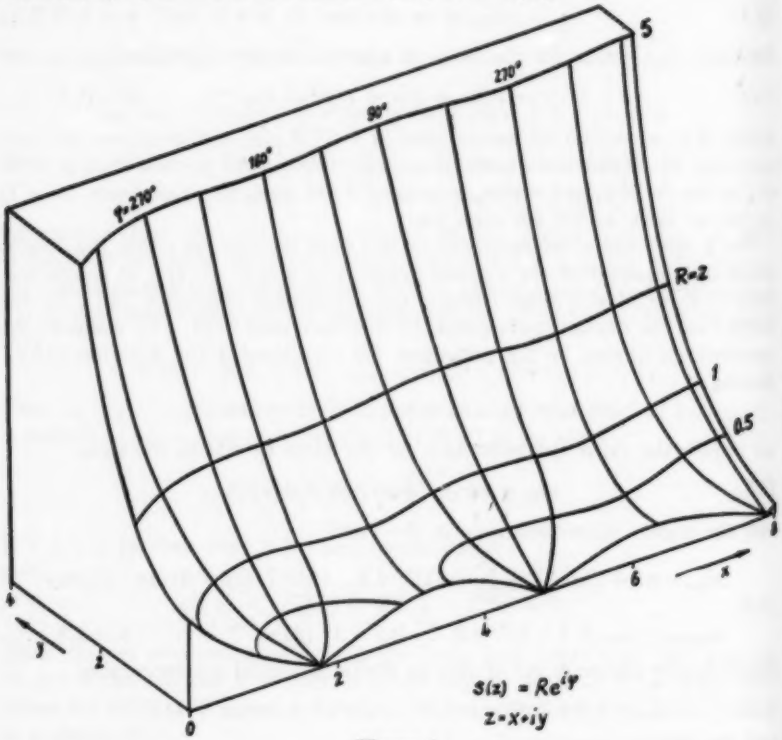

Figure 1

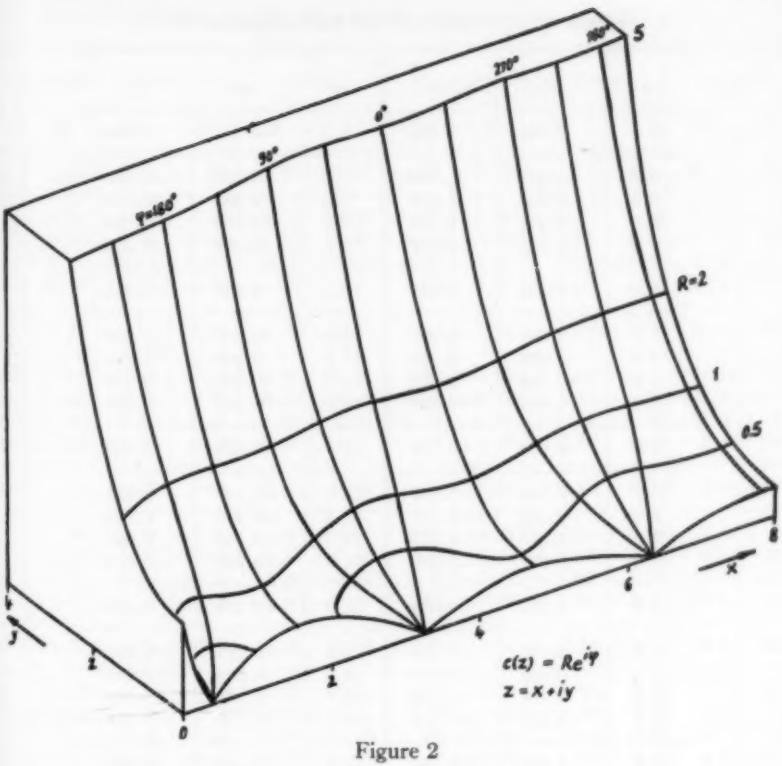

Figure 2

Zeros x_n of $c(z)$ and x_n^* of $s(z)$

n	x_n	x_n^*	n	x_n	x_n^*	n	x_n	x_n^*
0	0.41	1.77	10	31.43	33.00	20	62.84	64.41
1	3.27	4.81	11	34.57	36.14	21	65.98	67.55
2	6.36	7.92	12	37.71	39.28	22	69.12	70.69
3	9.48	11.04	13	40.85	42.42	23	72.26	73.83
4	12.61	14.17	14	43.99	45.56	24	75.41	76.98
5	15.74	17.31	15	47.14	48.71	25	78.55	80.12
6	18.88	20.44	16	50.28	51.85			
7	22.01	23.58	17	53.42	54.99			
8	25.15	26.72	18	56.56	58.13			
9	28.29	29.86	19	59.70	61.27			

Values of $c(z)$ and $s(z)$ for real argument $z = x$

x	$c(x)$	$s(x)$	x	$c(x)$	$s(x)$
0	1.253	1.253	7.5	−0.331	0.146
0.2	0.362	1.193	8.0	−0.350	−0.030
0.4	0.013	1.086	8.5	−0.283	−0.190
0.6	−0.241	0.951	9.0	−0.153	−0.294
0.8	−0.425	0.797	9.5	0.007	−0.323
1.0	−0.556	0.632	10.0	0.158	−0.272
1.2	−0.643	0.465	10.5	0.263	−0.159
1.4	−0.690	0.292	11.0	0.299	−0.012
1.6	−0.702	0.129	11.5	0.263	0.130
1.8	−0.682	−0.022	12.0	0.164	0.236
2.0	−0.635	−0.158	12.5	0.029	0.280
2.2	−0.566	−0.277	13.0	−0.107	0.255
2.4	−0.478	−0.376	13.5	−0.212	0.169
2.6	−0.377	−0.451	14.0	−0.263	0.045
2.8	−0.267	−0.503	14.5	−0.248	−0.085
3.0	−0.153	−0.531	15.0	−0.174	−0.190
3.2	−0.040	−0.536	15.5	−0.061	−0.247
3.4	0.069	−0.519	16.0	0.064	−0.241
3.6	0.168	−0.482	16.5	0.169	−0.178
3.8	0.257	−0.427	17.0	0.230	−0.074
4.0	0.330	−0.357	17.5	0.234	0.045
4.5	0.438	−0.142	18.0	0.181	0.150
5.0	0.430	0.085	18.5	0.085	0.215
5.5	0.319	0.271	19.0	−0.029	0.227
6.0	0.142	0.376	19.5	−0.133	0.183
6.5	−0.056	0.383	20.0	−0.202	0.096
7.0	−0.226	0.297			

Values of $c(z)$ for complex argument $z = x + iy$

x	$y=0$	$y=1$		$y=2$		$y=3$		$y=4$		$y=5$	
	$\Re$[1]	$\Re$	$\Im$	$\Re$	$\Im$	$\Re$	$\Im$	$\Re$	$\Im$	$\Re$	$\Im$
0	1.25	−0.31	−1.56	−1.71	−2.96	−4.54	−5.79	−11.1	−12.3	−26.5	−27.8
1	−0.56	−1.09	−0.46	−2.75	−0.66	−6.57	−0.58	−15.5	0.10	−37.0	2.40
2	−0.64	−0.93	0.39	−1.92	1.40	−4.12	4.12	− 9.01	11.1	−20.1	28.9
3	−0.15	−0.14	0.66	0.01	1.99	0.68	5.12	3.02	12.8	9.94	31.6
4	0.33	0.56	0.36	1.49	0.97	4.13	2.19	11.1	4.74	29.4	10.1
5	0.43	0.64	−0.16	1.52	−0.61	3.81	−1.95	9.56	− 5.86	23.9	−16.9
6	0.14	0.18	−0.46	0.31	−1.43	0.50	−3.92	0.59	−10.5	− 0.13	−27.5
7	−0.23	−0.38	−0.33	−0.99	−0.95	−2.80	−2.42	− 7.81	− 5.99	−21.5	−14.7
8	−0.35	−0.54	0.07	−1.30	0.27	−3.38	0.93	− 8.85	2.96	−23.0	9.07
9	−0.15	−0.22	0.36	−0.46	1.12	−1.03	3.12	− 2.33	8.51	− 5.10	23.0
10	0.16	0.26	0.31	0.68	0.92	1.95	2.44	5.55	6.33	15.6	16.3
11	0.30	0.46	−0.01	1.12	−0.06	2.99	−0.28	7.98	− 1.10	21.4	− 3.82
12	0.16	0.24	−0.29	0.54	−0.90	1.34	−2.54	3.33	− 6.99	8.23	−19.1
13	−0.11	−0.18	−0.29	−0.47	−0.89	−1.35	−2.41	− 3.88	− 6.37	−11.0	−16.8
14	−0.26	−0.41	−0.04	−1.00	−0.10	−2.67	−0.18	− 7.20	− 0.23	−19.4	0.02
15	−0.17	−0.26	0.23	−0.61	0.73	−1.55	2.06	− 4.00	5.71	−10.2	15.8
16	0.06	0.10	0.28	0.29	0.86	0.86	2.34	2.53	6.26	7.34	16.7
17	0.23	0.36	0.08	0.88	0.21	2.37	0.53	6.44	1.25	17.6	2.89
18	0.18	0.27	−0.18	0.65	−0.58	1.68	−1.65	4.43	− 4.59	11.7	−12.7
19	−0.03	−0.05	−0.26	−0.15	−0.82	−0.47	−2.24	− 1.41	− 6.05	− 4.25	−16.3
20	−0.20	−0.32	−0.11	−0.78	−0.31	−2.09	−0.81	− 5.71	− 2.06	−15.5	− 5.19

[1] $\Im = 0$.

Values of $s(z)$ for complex argument $z = x + iy$

x	$y=0$	$y=1$		$y=2$		$y=3$		$y=4$		$y=5$	
	$\Re$[1]	$\Re$	$\Im$	$\Re$	$\Im$	$\Re$	$\Im$	$\Re$	$\Im$	$\Re$	$\Im$
0	1.25	1.76	−0.51	3.02	−1.77	5.80	−4.55	12.3	−11.0	27.8	−26.5
1	0.63	0.68	−0.96	0.73	−2.72	0.69	−6.62	− 0.10	−15.5	− 2.41	−37.1
2	−0.16	−0.41	−0.71	−1.40	−1.84	−4.12	−4.10	−11.0	− 9.01	−28.8	−20.2
3	−0.53	−0.85	−0.06	−2.05	0.04	−5.15	0.70	−12.7	3.01	−31.5	9.94
4	−0.36	−0.51	0.46	−1.03	1.46	−2.21	4.12	− 4.75	11.1	−10.1	29.4
5	0.08	0.18	0.50	0.60	1.47	2.20	3.73	5.86	9.55	16.9	23.8
6	0.37	0.59	0.12	1.47	0.28	3.93	0.48	10.5	0.58	27.6	− 0.14
7	0.29	0.44	−0.30	0.99	−0.97	2.43	−2.78	5.99	− 7.80	14.9	−21.5
8	−0.03	−0.07	−0.41	−0.27	−1.24	−0.92	−3.35	− 2.95	− 8.81	− 9.05	−23.0
9	−0.30	−0.47	−0.15	−1.16	−0.43	−3.14	−1.02	− 8.53	− 2.32	−23.0	− 5.10
10	−0.28	−0.41	0.21	−0.96	0.67	−2.46	1.95	− 6.34	5.55	−16.2	15.6
11	−0.02	−0.01	0.35	0.05	1.09	0.28	1.09	1.09	7.99	3.81	21.3
12	0.23	0.37	0.18	0.93	0.52	2.54	1.34	6.99	3.33	19.2	8.24
13	0.25	0.39	−0.14	0.92	−0.46	2.52	−1.35	6.37	− 3.87	16.6	−11.0
14	0.04	0.05	−0.31	0.10	−0.96	0.20	−2.45	0.23	− 7.20	− 0.03	−19.4
15	−0.19	−0.30	−0.19	−0.76	−0.58	−2.07	−1.54	− 5.72	− 3.99	−15.7	−10.3
16	−0.24	−0.37	0.09	−0.89	0.29	−2.35	0.86	− 6.27	2.53	−16.7	7.34
17	−0.08	−0.11	0.27	−0.23	0.85	−0.54	2.36	− 1.25	6.44	− 2.89	17.5
18	0.15	0.23	0.21	0.60	0.62	1.65	1.68	4.59	4.43	12.8	11.7
19	0.22	0.35	−0.04	0.84	−0.15	2.25	−0.46	6.06	− 1.41	16.3	− 4.23
20	0.09	0.14	−0.24	0.32	−0.75	0.81	−2.08	2.06	− 5.70	5.19	−15.6

[1] $\Im = 0$.

References

1. E. Kreyszig, *On the zeros of the Fresnel integrals*, Can. J. Math., *9* (1957), 118–131.
2. ———, *Ueber den allgemeinen Integralsinus Si(z,α)*, Acta math., *85* (1951), 117–181.
3. ———, *Der allgemeine Integralkosinus Ci(z,α)*, Acta math., *89* (1953), 107–131.
4. E. Lommel, *Beugungserscheinungen einer kreisrunden Oeffnung und eines kreisrunden Schirmchens*, Abh. Bayr. Akad. Wiss. Muenchen *15* (1886), 231–330.

University of Ottawa
and
Ohio State University, Columbus, Ohio

ON A METRIC THAT CHARACTERIZES DIMENSION

J. DE GROOT

1. Introduction. Sometimes it is possible to characterize topological properties of a metrizable space M by claiming that a certain (topology-preserving) metric ρ can be introduced in M. For example:

(α) A metrizable space C is compact, that is, is a compactum, if and only if C is totally bounded[1] in every metric.

(β) A metrizable space M is separable, if and only if there exists a totally bounded metric in M.

(γ) A (non-empty) metrizable space M is 0-dimensional (dim $M = 0$), if and only if there exists a metric ρ in M which satisfies—instead of the triangle axiom—the stronger axiom

$$\rho(y, z) \leqslant \max[\rho(x, y), \rho(x, z)], \tag{1.1}$$

(that is, every "triangle" in this metric has two equal "sides" and the third "side" is smaller than or equal to the other ones) (see **2**, **3**).

Nagata **(7)** gave a characterization of a metrizable space M of dim $\leqslant n$ (for every non-negative integer n) by means of a certain metric, which he showed to be equivalent with (γ) in the case $n = 0$. However, this characterization (see §**2**) is rather complicated. In this note we give another generalization of (γ) which gives a simplification of Nagata's result for arbitrary dimension n, but only for the case of *separable* metrizable spaces, i.e., metrizable spaces with a countable base.

THEOREM. *A topological space M is a separable metrizable space of dimension $\leqslant n$ if and only if one can introduce a totally bounded metric ρ in M satisfying the following condition: for every $n + 3$ points*

$$x, y_1, y_2, y_3, \ldots y_k, \ldots y_{n+2}$$

in M there is a triplet of indices i, j, k, such that

$$\rho(y_i, y_j) \leqslant \rho(x, y_k), \qquad (i \neq j). \tag{1.2}$$

COROLLARY. *A compactum has dimension $\leqslant n$, if and only if one can introduce a metric ρ, such that for every $n + 3$ points x, $y_k (k = 1,2, \ldots, n + 2)$ the relation* 1.2 *holds for suitable i, j, k.*

Received May 28, 1957.

[1] ϵ-net: A finite number of points p such that the system of ϵ-neighbourhoods cover the space. Totally bounded: there is an ϵ-net for every $\epsilon > 0$. See **(1)** in general for our terminology. See **(4)** for dimension theory in separable metrizable spaces and **(5; 6)** for dimension theory in metrizable spaces.

It has to be observed that condition 1.2 is essentially weaker than the condition which is satisfied by Nagata's metric **(7)** (see also § 2). Indeed, the ordinary metric of a segment of real numbers is a metric ρ with 1.2 (for the case $n = 2$), but does not satisfy Nagata's condition.

2. Proof of Theorem. Suppose M is a separable metric space with $\dim M \leqslant n$. Since M is separable, we can embed M, according to a theorem of Hurewicz, in a compactum $\bar{M}$, such that M is dense in $\bar{M}$, and

$$\dim M = \dim \bar{M} \leqslant n.$$

We introduce in $\bar{M}$ the metric ρ of Nagata **(7)**, which has the following characterizing property: for every $\epsilon > 0$ and for every point $x \in \bar{M}$ the relations[2]

2.1 $$\rho(U_\delta(x), y_k) < \epsilon \qquad (k = 1, 2, \ldots n+2),$$

where $U_\delta(x)$ is the set of all points p with $\rho(x, p) < \delta$, imply

2.2 $$\min_{i \neq j} \rho(y_i, y_j) < \epsilon.$$

It is easy to see that this metric ρ in particular satisfies our condition 1.2. Indeed, being given the points x, y_k $(k = 1,2, \ldots n+2)$, consider all ϵ with

$$\epsilon > \mu = \max_k \rho(x, y_k).$$

For these ϵ, 2.1 obviously holds, so 2.2 holds.
Since $\inf \epsilon = \mu$, we have

$$\min_{i \neq j} \rho(y_i, y_j) \leqslant \mu \qquad \text{q.e.d.}$$

Moreover, the metric ρ in the compact space $\bar{M}$ is necessarily totally bounded. Hence the metric ρ of $M \subset \bar{M}$ is also totally bounded and satisfies 1.2, which we had to prove.

Conversely, let M have a totally bounded metric satisfying 1.2. M is clearly separable. We shall now prove that $\dim M \leqslant n$.

M can be extended, just as every metric space, to a complete metric space $\bar{M}$ in which M is dense. Every sequence in M has a Cauchy sequence (fundamental sequence) as subsequence, since M is totally bounded under ρ. This Cauchy sequence converges in the complete $\bar{M}$. Hence $\bar{M}$ is compact and totally bounded under ρ, where ρ now denotes the natural extension of ρ (on M) to $\bar{M}$. Property 1.2 also holds in this extended metric ρ on $\bar{M}$. Indeed, suppose it does not hold for a set of certain points $\bar{x}$, $\bar{y}_k$. Then, since the distance function is continuous, we can determine small neighbourhoods of these points such that 1.2 does not hold for any set of points x, y_k chosen in these neighbourhoods respectively. We can, however, choose these points x, y_k from M, which leads to a contradiction. We shall now prove $\dim \bar{M} \leqslant n$, from which follows $\dim M \leqslant n$.

[2] The distance of the sets A and B is denoted by $\rho(A, B)$.

Consider an arbitrary finite open covering of $\bar{M}$. We have to find—according to the Lebesgue definition of dimension—a refinement of this covering of order $\leqslant n$ (i.e. each point of the refined covering is contained in at most $n+1$ elements of it).

Let $\sigma = 2\epsilon$ be a Lebesgue number of the given finite covering of $\bar{M}$. Choose a maximal set $p_1, p_2, \ldots, p_s$ in $\bar{M}$ such that $\rho(p_i, p_j) \geqslant \epsilon$ for all i, j with $i \neq j$. This set of points $\{p_i\}$ is an ϵ-net of $\bar{M}$ and the covering

$$\{U_\epsilon(p_i)\} \qquad (i = 1, 2, \ldots s) \tag{2.3}$$

is a refinement of the given covering. If a point $x \in \bar{M}$ belongs to at least $n+2$ elements of 2.3, we have $\rho(x, p_i) < \epsilon$ for $n+2$ different points p_i. Hence, using 1.2, $\rho(p_i, p_j) < \epsilon$ for suitable i, j with $i \neq j$, which is contradictory to the definition of $\{p_i\}$. Hence, the order of 2.3 is $\leqslant n$, so $\dim \bar{M} \leqslant n$.

3. Questions. The corollary admits an immediate generalization to semicompact[3] metrizable spaces, since we can apply in this case the sum theorem of dimension theory (a metric space which is the countable sum of closed subsets of dimension $\leqslant n$, has dimension $\leqslant n$), while the proof in the other direction is covered by Nagata's theorem, as mentioned in §2. So, our characterization by means of a metric satisfying 1.2 includes for example n-dimensional Euclidean spaces as well.

However, it remains uncertain whether in separable metric spaces M the property $\dim \leqslant n$ can be characterized by a metric satisfying 1.2 only. There might be a possibility that the condition of total boundedness can be omitted in this case, if the condition 1.2 is strengthened in the following way: there is a metric ρ in M which satisfies 1.2 and also, if $\rho(x,y_1) = \rho(x,y_2) = \ldots = \rho(x,y_{n+2})$,

$$\rho(y_i, y_j) < \rho(x, y_k), \quad \text{for suitable } i, j, k \qquad (i \neq j). \tag{3.1}$$

However, does there exist such a metric? For $n = 0$, the answer is in the affirmative **(4, §2)**.

The problem of generalizing the Theorem to metric spaces in general remains unanswered too.

---

[3]A space is semicompact if it is the sum of a countable number of compact spaces. Every locally compact, separable, metrizable space is semicompact, since such a space can be compactified by one point.

References

1. P. Alexandrov and H. Hopf, *Topologie* (Berlin 1935).
2. J. de Groot and H. de Vries, *A note on non-archimedean metrizations*, Indagationes Math., *17* (1955), 222–224.
3. J. de Groot, *Non-archimedian metrics in topology*, Proc. Amer. Math. Soc., *7* (1956), 948–953.
4. W. Hurewicz, H. Wallman, *Dimension Theory* (Princeton 1941).
5. M. Katětov, *On the dimension of non-separable spaces I*, Tsjechoslov. Mat. Zj., *2* (77) (1952), 333–368.
6. K. Morita, *Normal families and dimension theory for metric spaces*, Math. Ann., *128* (1954), 350–362.
7. J. Nagata, *On a relation between dimension and metrization*, Proc. Jap. Ac., *32* (1956), 237–240.

Mathematisch Instituut
University of Amsterdam

GRAPHS WITH GIVEN GROUP AND GIVEN GRAPH-THEORETICAL PROPERTIES

GERT SABIDUSSI

1. Introduction. In 1938 Fruchț **(2)** proved the following theorem:

(1.1). THEOREM. *Given any finite group G there exist infinitely many non-isomorphic connected graphs X whose automorphism group is isomorphic to G.*

Later, the same author showed **(3)** that this theorem still holds, if the words "connected graphs X" are replaced by "connected regular graphs X of degree 3." There is, of course, no reason to assume that such graphs play any distinguished rôle, and that similar theorems do not hold for degrees > 3. Indeed it can be shown that (1.1) holds with "connected graphs X" replaced by "connected regular graphs X of degree n, where n is any integer $\geqslant 3$."

It is only natural, then, to investigate whether the property that a graph X be regular of degree n is the only graph-theoretical property of X which can be prescribed together with the automorphism group. Consider the following properties $P_j (j = 1,2,3,4)$ of X:

P_1: The connectivity **(6)** of X is n, where n is an integer $\geqslant 1$.

P_2: The chromatic number **(1)** of X is n, where n is an integer $\geqslant 2$.

P_3: X is regular of degree n, where n is an integer $\geqslant 3$.

P_4: X is spanned by a graph $\bar{Y}$ homeomorphic to a given connected graph Y.

Call a graph X *fixed-point-free* if there is no vertex x of X which is invariant under all automorphisms of X.

The following theorem contains the main results of this paper:

(1.2) THEOREM. *Given a finite group G of order > 1 and an integer j, $1 \leqslant j \leqslant 4$, there exist infinitely many non-homeomorphic connected fixed-point-free graphs X such that* (i) *the automorphism group of X is isomorphic to G, and* (ii) *X has property P_j.*

The principal tool in deriving these results is the graph multiplication "$\times$" defined in **(5)**. A typical proof of the statements of (1.2) runs as follows:

(a) Construct a connected fixed-point-free prime graph X' (for a definition of "prime" **(5**, (1.3)**)**) whose automorphism group is isomorphic to G.

(b) Construct a connected prime graph $X'' \ncong X'$ with trivial automorphism

Received April 29, 1956; in revised form October 5, 1956. Written with the partial support of the National Science Foundation Grant to Tulane University.

group and certain graph theoretical properties P_j' which are such that the product $X' \times X''$ has property P_j.

(c) Apply **(5**, Theorem (3.2)**)**, with the result that

The automorphism group of $X' \times X''$ is isomorphic to the automorphism group of X', that is, isomorphic to G.

By a *graph* X we mean an ordered triple $X = (V, E, f)$, where V and E are two disjoint sets (the sets of vertices and edges of X), and f is a function of E into the set V^* of unordered pairs of distinct elements of V such that if $e^* \in V^*$ there is *at most one* $e \in E$ with $fe = e^*$. To indicate that V and E are the sets of vertices and edges of a graph $X = (V, E, f)$ we shall write $V = V(X)$, $E = E(X)$. Edges will be written as unordered pairs of vertices (indicated by brackets). To describe a graph X it clearly suffices to give the set $V(X)$ and a certain set $E(X)$ of unordered pairs of elements of $V(X)$. All graphs considered in this paper are *finite*.

Let X be a graph. By $G(X)$ we denote the *automorphism group* of X. We can consider $G(X)$ as a group of one-one mappings of $V(X)$ onto itself.

2. Definition and properties of the graph product.

(2.1) DEFINITION: Let X, Y be graphs. By the *product* $X \times Y$ of X and Y is meant the following graph Z:

$$V(Z) = V(X) \times V(Y);$$

$[(x,y), (x',y')]$, where $x, x' \in V(X)$, $y, y' \in V(Y)$, is an edge of Z if $x = x'$ and $[y,y'] \in E(Y)$, or $y = y'$ and $[x,x'] \in E(X)$.

If we identify isomorphic graphs the multiplication thus defined is clearly *associative* and *commutative*. It has a *unit*, viz. the graph consisting of a single vertex and no edge.

(2.2) LEMMA. *The product of connected graphs is connected. The product of any graph by a disconnected graph is disconnected.*

(2.3) LEMMA. *If X is m-ply connected, and Y is n-ply connected, then $X \times Y$ is $(m + n)$-ply connected.*

Proof. We shall use a theorem of Whitney **(6**, Theorem 7**)**. X is m-ply connected implies: Given any pair of distinct vertices x, x' of X there exist m paths X_j of X such that

$$V(X_j) \cap V(X_k) = \{x, x'\}, \qquad j \neq k.$$

Y is n-ply connected means: Given any pair of distinct vertices y, y' of Y there exist n paths Y_j of Y such that

$$V(Y_j) \cap V(Y_k) = \{y, y'\}, \qquad j \neq k.$$

To show: Given any pair of distinct vertices (x,y), (x',y') of $Z = X \times Y$ there exist $m + n$ paths Z_j of Z such that

$$V(Z_j) \cap V(Z_k) = \{(x,y), (x',y')\}, \qquad j \neq k.$$

We have to consider two cases.

Case (1). Given (x,y), $(x',y') \in V(Z)$, where $x \neq x'$, $y \neq y'$. At most one of the paths X_j (Y_j) consists of a single edge. In that case let the notation be so chosen that X_m (Y_n) is that path. Let

$$V(X_j) = \{x, x_2^{(j)}, x_3^{(j)}, \ldots, x'\}, \qquad j \leqslant m;$$
$$V(Y_j) = \{y, y_2^{(j)}, y_3^{(j)}, \ldots, y'\}, \qquad j \leqslant n.$$

Define paths Z_j, Z_{m+k} of Z as follows:

$$V(Z_j) = \{(x,y), (x_2^{(j)},y), (x_2^{(j)},y_2^{(n)}), (x_2^{(j)},y_3^{(n)}), \ldots, (x_2^{(j)},y'), (x_3^{(j)},y'), \ldots, (x',y')\}, \qquad j \leqslant m-1;$$
$$V(Z_m) = \{(x,y), (x_2^{(m)},y), \ldots, (x',y), (x',y_2^{(n)}), \ldots, (x',y')\};$$
$$V(Z_{m+k}) = \{(x,y), (x,y_2^{(k)}), (x_2^{(m)},y_2^{(k)}), (x_3^{(m)},y_2^{(k)}), \ldots, (x',y_2^{(k)}), (x'y_3^{(k)}), \ldots, (x',y')\}, \qquad k \leqslant n-1;$$
$$V(Z_{m+n}) = \{(x,y), (x,y_2^{(n)}), \ldots, (x,y'), (x_2^{(m)},y'), \ldots, (x',y')\}.$$

Case (2). Given (x,y), $(x',y) \in V(Z)$, where $x \neq x'$. Let y' be any vertex of Y distinct from y. Using the same notation as in case (1), define Z_j, Z_{m+k} as follows:

$$V(Z_j) = \{(x,y), (x_2^{(j)},y), \ldots, (x',y)\}, \qquad j \leqslant m;$$
$$V(Z_{m+k}) = \{(x,y), (x,y_2^{(k)}), (x_2^{(m)},y_2^{(k)}), (x_3^{(m)},y_2^{(k)}), \ldots, (x',y_2^{(k)}), (x',y)\}, \qquad k \leqslant n.$$

In both cases the $m + n$ paths Z_j, Z_{m+k} of Z have the required properties.

(2.4) LEMMA. *Let X, Y be graphs of connectivity m and n respectively. If there is an $x \in V(X)$ of degree m, and a $y \in V(Y)$ of degree n, then the connectivity of $X \times Y$ is $m + n$.*

Proof. Let V_x, V_y, $V_{(x,y)}$ be the sets of those vertices of X, Y and $X \times Y$ which are joined with $x \in V(X)$, $y \in V(Y)$, and $(x,y) \in V(X \times Y)$ respectively. Then the definition of the graph product implies that

$$V_{(x,y)} = (V_x \times \{y\}) \cup (\{x\} \times V_y).$$

Hence the degree of (x,y) in $X \times Y$ is $m + n$. It follows that every subgraph I of $X \times Y$ with $V(I) = V_{(x,y)}$ is an isthmoid[1] of order $m + n$ of $X \times Y$ (with one component of $(X \times Y) - I$ consisting of the vertex (x,y) alone). Hence the connectivity of $X \times Y$ is $\leqslant m + n$. By (2.3) the connectivity of $X \times Y$ is $\geqslant m + n$, and this proves Lemma (2.4).

[1]An *isthmoid* of a graph X is a subgraph I of X such that $X - I$ is disconnected. $X - I$ is the maximal subgraph X' of X with $V(X') = V(X) - V(I)$.

(2.5) LEMMA. *If X and Y are regular of degree m and n respectively, then $X \times Y$ is regular of degree $m + n$.*

The proof of this Lemma is contained in the proof of (2.4).

(2.6) LEMMA. *Let $\chi(X)$, $\chi(Y)$, $\chi(X \times Y)$ be the chromatic numbers of X, Y and $X \times Y$ respectively. Then $\chi(X \times Y) = \max(\chi(X), \chi(Y))$.*

Proof. The maximal subgraphs X_y, Y_x of $X \times Y$ with

$$V(X_y) = V(X) \times \{y\},\ y \in V(Y),$$
$$V(Y_x) = \{x\} \times V(Y),\ x \in V(X),$$

are isomorphic to X and Y respectively. Hence $\chi(X \times Y) \geqslant m$, where $m = \max(\chi(X), \chi(Y))$. Let c_X, c_Y be m-colorings of X and Y respectively (an m-coloring of X is a function c_X of $V(X)$ into J_m, the group of integers (mod m), such that $[x,x'] \in E(X)$ implies $c_X(x) \neq c_X(x')$; likewise for Y). Define a function c of $V(X \times Y)$ into J_m by

$$c(x,y) = c_X(x) + c_Y(y),\ x \in V(X),\ y \in V(Y).$$

c is an m-coloring of $X \times Y$. To show:

$$[(x,y),(x',y')] \in E(X \times Y) \rightarrow c(x,y) \neq c(x',y');$$

$$[(x,y),(x',y')] \in E(X \times Y) \rightarrow x = x',\ [y,y'] \in E(Y), \text{ or } y = y',\ [x,x'] \in E(X).$$

It suffices to consider the first case: $x = x' \rightarrow c_X(x) = c_X(x')$; $[y,y'] \in E(Y) \rightarrow c_Y(y) \neq c_Y(y')$. Hence

$$c(x,y) = c_X(x) + c_Y(y) \neq c_X(x') + c_Y(y') = c(x',y').$$

Since c is an m-coloring of $X \times Y$, it follows that $\chi(X \times Y) \leqslant m$.

(2.7) LEMMA. *Let X, Y, Z be connected graphs such that* (i) $\alpha_0(Z) \geqslant 2\alpha_0(X) - 2$ (α_0 = *number of vertices*); (ii) *Z contains a Hamiltonian circuit H*; (iii) *Z is spanned by a graph $\bar{Y}$ homeomorphic to Y with $E(\bar{Y}) \cap E(H) \neq \square$ (= the empty set). Then $X \times Z$ is spanned by a graph $\bar{Y}$ homeomorphic to Y.*

Proof. Let X' be a (connected) spanning tree of X, and let $E(X') = \{e_1, \ldots, e_{m-1}\}$, $m = \alpha_0(X)$. For each $x_i \in V(X')$ define $E_i = \{k | x_i$ is incident with e_k, $e_k \in E(X')\}$. Since X' is a tree, we can assume that $x_1 \in V(X')$ is of degree 1 in X'. Let

$$V(Z) = V(\bar{Y}) = \{z_1, \ldots, z_n\},\ n \geqslant 2m - 2,$$

and let the notation be so chosen that

(1) $E(H) = \{[z_1,z_2], [z_2,z_3], \ldots, [z_{n-1},z_n], [z_n,z_1]\}$, and
(2) $[z_1,z_2] \in E(\bar{Y}) \cap E(H)$. Let H_i, $\bar{Y}_i$ be given by

$$V(H_i) = V(\bar{Y}_i) = \{x_i\} \times V(\bar{Y}),$$
$$E(H_i) = \{[(x_i,z_j),(x_i,z_k)] | [z_j,z_k] \in E(H)\},$$
$$E(\bar{Y}_i) = \{[(x_i,z_j),(x_i,z_k)] | [z_j,z_k] \in E(\bar{Y})\}.$$

Notice that

$$\bigcup_{i=1}^{m} V(H_i) = V(X \times Z).$$

Consider the following subgraph P of $X \times Z$:

$$V(P) = \bigcup_{i=2}^{m} V(H_i) \cup \{(x_1,z_1), (x_1,z_2)\},$$

$$E(P) = \bigcup_{i=2}^{m} (E(H_i) - \{[(x_i,z_{2k-1}), (x_i,z_{2k})] | k \in E_i\}) \cup$$

$$\bigcup_{k=1}^{m-1} \{[(x^{(k)},z_{2k-1}), (y^{(k)},z_{2k-1})], [(x^{(k)},z_{2k}), (y^{(k)},z_{2k})]\},$$

where $[x^{(k)},y^{(k)}] = e_k (k = 1, \ldots, m-1)$. It can be easily checked that (1) P is connected, (2) the degree of (x_1,z_1) and (x_1,z_2) in P is 1, (3) the degree in P of any other vertex of P is 2. Hence P is a path joining (x_1,z_1) and (x_1,z_2), and containing all vertices of $(X \times Z) - \hat{Y}_1$. Now let $\hat{Y}$ be given by

$$V(\hat{Y}) = V(X \times Z),$$
$$E(\hat{Y}) = (E(\hat{Y}_1) - \{[(x_1,z_1), (x_1,z_2)]\}) \cup E(P).$$

Then clearly $\hat{Y}$ spans $X \times Z$, and is homeomorphic to Y.

(2.8) LEMMA. *Every connected graph X containing a vertex or an edge which is not contained in a 4-circuit of X is prime.*

Proof. Suppose $X = Y \times Z$, where $\alpha_0(Y), \alpha_0(Z) \geqslant 2$. Let

$$(y,z) \in V(X), y \in V(Y), z \in V(Z).$$

Since X is connected, both Y and Z are connected; hence by (2.5) the degree of y in Y and the degree of z in Z must be $\geqslant 1$. Let y', z' be vertices joined with y and z in Y and Z respectively. Then the subgraph C of X given by

$$V(C) = \{(y,z), (y,z'), (y',z'), (y',z)\},$$
$$E(C) = \{[(y,z), (y,z')], [(y,z'), (y',z')], [(y',z'), (y',z)], [(y',z), (y,z)]\}$$

is a 4-circuit of X containing (y,z). The same proof applies to edges.

(2.9) LEMMA. *The product of a fixed-point-free graph X by any graph Y is fixed-point-free.*

Proof. Let $x \in V(X)$. Since X is fixed-point-free, there is a $\phi \in G(X)$ such that $\phi x \neq x$. Then the function ϕ^* given by $\phi^*(x,y) = (\phi x,y)$ is an automorphism of $X \times Y$, and $\phi^*(x,y) \neq (x,y)$ for all $y \in V(Y)$. Hence $X \times Y$ is fixed-point-free.

(2.10) LEMMA. **(5, (3.2))**. *If X and Y are relatively prime, then $G(X \times Y) \cong G(X) \times G(Y)$.*

For a definition of "relatively prime" cf. **(5, (1.3))**.

3. Existence of graphs with given group and given graph theoretical properties. We shall now prove the four theorems stated as Theorem (1.2). It should be emphasized that the constructions given in this paragraph are by no means the only possible ones. They have been chosen mainly to demonstrate the usefulness of graph multiplication.

(3.1) *Definition:* Let X be a graph without isolated vertices. By $\bar{X}$ we mean the graph defined by

(1) $V(\bar{X}) = \{(x,e) \in V(X) \times E(X) | x \text{ is incident with } e\}$;

(2) given (x,e), $(x',e') \in V(\bar{X})$, then $[(x,e), (x',e')] \in E(\bar{X})$ if and only if $x = x'$, $e \neq e'$, or $x \neq x'$, $e = e'$.

The following properties of $\bar{X}$ are obvious from the definition.

(3.2) LEMMA. *Let X be as in* (3.1). (i) *If X is connected or cyclically connected, then so also is $\bar{X}$.* (ii) *If X is regular of degree $n \geqslant 1$, then $\bar{X}$ is likewise of degree n.* (iii) *If no component of X is a circuit, then X and $\bar{X}$ are not homeomorphic. If X is an n-circuit, then $\bar{X}$ is a $2n$-circuit.* (iv) *If X is connected, then $\bar{X}$ is prime.*

(3.3) LEMMA. *Let X be as in* (3.1). *If X is fixed-point-free and without fixed edge*[3], *then so also is $\bar{X}$. If no component of X is a circuit, then $G(X) \cong G(\bar{X})$.*

Proof. Given $\phi \in G(X)$ define $\bar{\phi}$: $V(\bar{X}) \to V(\bar{X})$ by $\bar{\phi}(x,e) = (\phi x, \phi e)$. Then clearly $\bar{\phi} \in G(\bar{X})$, and $\phi \to \bar{\phi}$ is an isomorphism of $G(X)$ into $G(\bar{X})$.

Define an equivalence relation $\sim$ on $V(\bar{X})$ by $(x,e) \sim (x',e')$ if and only if $x = x'$. Let $\mathbf{X}$ be the graph given by

(i) $V(\mathbf{X}) = V(\bar{X})/\sim$;

(ii) $[\mathfrak{x},\mathfrak{x}'] \in E(\mathbf{X})$, where $\mathfrak{x}, \mathfrak{x}' \in V(\mathbf{X})$, if and only if there exist $(x, e) \in \mathfrak{x}$ and $(x',e') \in \mathfrak{x}'$ such that $[(x,e), (x',e')] \in E(\bar{X})$.

Then clearly $\mathbf{X} \cong X$. By p denote the natural projection of $V(\bar{X})$ onto $V(\mathbf{X})$.

$G(\bar{X})$ preserves the relation $\sim$. Let $(x_1,e_1) \sim (x_2,e_2)$, so that $x_1 = x_2$, and let $\bar{\psi} \in G(\bar{X})$. Put $\bar{\psi}(x_i,e_i) = (x_i',e_i')$ $(i = 1,2)$. To show that $x_1' = x_2'$. $(x_1,e_1) \sim (x_2,e_2) \to (x_1,e_1) = (x_2,e_2)$ or $\epsilon = [(x_1,e_1), (x_2,e_2)] \in E(\bar{X})$. Hence $(x_1',e_1') = (x_2',e_2')$, and hence $x_1' = x_2'$, or $\epsilon' = [(x_1',e_1'), (x_2',e_2')] \in E(\bar{X})$. In the latter case either

(1) $x_1' = x_2'$, $e_1' \neq e_2'$, or

(2) $x_1' \neq x_2'$, $e_1' = e_2'$.

We have to show that (2) leads to a contradiction. Assume (2). It is easily seen that then there is no 3-circuit of $\bar{X}$ containing ϵ'. Hence there is no 3-circuit of $\bar{X}$ containing ϵ. Therefore $x_1 = x_2$ is of degree 2 in X, which in turn implies that (x_i,e_i) $(i = 1,2)$ are of degree 2 in $\bar{X}$. Since no component of X is a circuit, no component of $\bar{X}$ is a circuit (cf. (3.2) (ii), (iii)). Hence $\bar{X}$ contains a vertex

$$(y_1,e_{i_1})$$

[3] An edge e of X is *fixed*, if $\phi e = e$ for all $\phi \in G(X)$.

and a path $\bar{P}$ with

$$V(\bar{P}) = \{(y_1,e_{i_1}), \ldots, (y_n,e_{i_n})\}, \quad E(\bar{P}) = \{\epsilon_1, \ldots, \epsilon_{n-1}\},$$

such that

(α) $$(y_{n-1},e_{i_{n-1}}) = (x_1,e_1), \quad (y_n,e_{i_n}) = (x_2,e_2);$$

(β) $$(y_1,e_{i_1})$$

is of degree $\neq 2$ in $\bar{X}$;

(γ) $$(y_k,e_{i_k})$$

is of degree 2 in $\bar{X}$ for all $k \neq 1$.

We show that

$$n = 2m + 1, \; e_{i_{2k-1}} = e_{i_{2k}}, \; y_{2k} = y_{2k+1}, \qquad k < m.$$

For the proof notice that $(y,e) \in V(\bar{X})$ and $y \in V(X)$ are always of the same degree. $\epsilon_1 \in E(\bar{X})$ implies

(a) $$y_1 = y_2, \; e_{i_1} \neq e_{i_2},$$

or

(b) $$y_1 \neq y_2, \; e_{i_1} = e_{i_2}.$$

(a) is impossible because

$$(y_1,e_{i_1}) \text{ and } (y_2,e_{i_2}),$$

and hence y_1 and y_2, have different degrees. Hence (b) must hold. $\epsilon_2 \in E(\bar{X})$ implies

(c) $$y_2 = y_3, \; e_{i_2} \neq e_{i_3},$$

or

(d) $$y_2 \neq y_3, \; e_{i_2} = e_{i_3}.$$

Suppose (d) holds. Then (b) and (d) imply

$$e_{i_1} = e_{i_2} = e_{i_3},$$

which is incident with y_1, y_2, y_3. Two of these vertices must be equal: (b) and (d) imply $y_1 = y_3$. But then

$$(y_1,e_{i_1}) = (y_3,e_{i_3}),$$

so that

$$(y_3,e_{i_3})$$

is of degree 1, a contradiction. Hence (c) must hold. The rest of the assertion follows in a similar way by induction. We shall express the fact that $n = 2m + 1$ by saying that the "distance" of ϵ from

$$(y_1,e_{i_1})$$

is odd. Since ϵ and ϵ' are similar under $\bar{\psi}$, there is a vertex

$$(z_1,e_{j_1})$$

and a path $\bar{Q}$, similar under $\bar{\psi}$ to

(y_1,e_{i_1})

and $\bar{P}$ respectively, and such that (α), (β), (γ) are satisfied with respect to ϵ'. By the same argument as above it then follows that the distance of ϵ' from

(x_1,e_{j_1})

is even. But this contradicts the similarity of ϵ and ϵ'.

Given $\bar{\psi} \in G(\bar{X})$ define

$$\psi: V(\mathbf{X}) \to V(\mathbf{X}) \text{ by } \psi x = p\bar{\psi}(x,e),$$

where $(x,e) \in p^{-1}\mathbf{x}$, $\mathbf{x} \in V(\mathbf{X})$. Since $\bar{\psi}$ preserves equivalence, ψ is in $G(\mathbf{X})$, and $h: G(\bar{X}) \to G(\mathbf{X})$ given by $h\bar{\psi} = \psi$ is a homomorphism. Consider

$$\text{Ker } h = \{\bar{\psi} | \bar{\psi}(x,e) \sim (x,e)\}.$$

Let $e = [x,y] \in E(X)$. Then $[(x,e), (y,e)] \in E(\bar{X})$. For $\bar{\psi} \in \text{Ker } h$ put

$$\bar{\psi}(x,e) = (x_1,e_1), \quad \bar{\psi}(y,e) = (y_1,e_2).$$

Then $x = x_1$, $y = y_1$, and $[(x_1,e_1), (y_1,e_2)] \in E(\bar{X})$. Hence

(1) $x_1 = y_1$, $e_1 \neq e_2$, or

(2) $x_1 \neq y_1$, $e_1 = e_2 = [x_1,y_1]$.

(1) is impossible since it implies $x = y$; (2) implies $e_1 = e_2 = e$, so that $\bar{\psi}(x,e) = (x,e)$. Hence Ker $h = 1$, and h is an isomorphism.

The assertion about fixed vertices and edges follows from the fact that $\bar{\phi}$ given by $\bar{\phi}(x,e) = (\phi x,\phi e)$ is in $G(\bar{X})$.

All constructions in this paragraph are based on the following theorem:

(3.4) THEOREM. *Given a finite group G of order > 1, there exist infinitely many non-homeomorphic cyclically connected fixed-point-free prime graphs X_i containing no fixed edge, and such that $G(X_i) \cong G$.*

Proof. By **(3**, Theorem 4.1**)** there exists at least one such graph, X_1. By induction, let $X_{i+1} = \bar{X}_i$, $i \geq 1$. Then by (3.2) and (3.3) all X_i have the required properties. Since X_1 is regular of degree 3, no X_i is a circuit.

(3.5) THEOREM. *Given a finite group G of order > 1 and a positive integer n, there exist infinitely many non-homeomorphic fixed-point-free graphs X of connectivity n whose automorphism group is isomorphic to G.*

Proof. Given any graph X denote the connectivity of X by $c(X)$. For $n = 1$, (3.5) has been proved in **(2, §2)**. We can therefore assume that $n \geq 2$.

Case (1). $n = 2$. Let X' be a graph with the properties stated in (3.4). In particular, $c(X') \geq 2$. By subdividing each edge e of X' by a vertex x_e we obtain a graph X with $c(X) = 2$. X is prime, since no circuit of X is of order < 6 (cf. (2.8)). Since X' is not a circuit, $G(X) \cong G(X') \cong G$. X is fixed-point-free because X' is fixed-point-free and contains no fixed edge.

Case (2): $n \geq 3$. Let Y_k, $k \geq 1$, be the graph given by

$$V(Y_k) = \{0,1,\ldots,k+5\},$$
$$E(Y_k) = \{[0,1],\ [0,2],\ [2,3],\ [0,4],\ [4,5],\ldots,[k+4,\ k+5]\}.$$

Then (i) 1 is a vertex of degree 1 of Y_k; (ii) $c(Y_k) = 1$; (iii) $G(Y_k) = 1$; (iv) Y_k and $Y_{k'}$ are relatively prime if $k \neq k'$. It follows from (2.4) that

$$Y^{(m)} = Y_1 \times \ldots \times Y_m$$

is a graph of connectivity m, and from (2.10) that $G(Y^{(m)}) = 1$. By (2.5), $(1,\ldots,1)$ is a vertex of degree m of $Y^{(m)}$.

Let X be as in case (1). Then X and $Y^{(n-2)}$, $n \geq 3$, are relatively prime, and satisfy the hypotheses of (2,4). Hence $c(X \times Y^{(n-2)}) = n$, and by (2.10),

$$G(X \times Y^{(n-2)}) \cong G(X) \times G(Y^{(n-2)}) \cong G.$$

By (2.9), $X \times Y^{(n-2)}$ is fixed-point-free.

(3.6) THEOREM. *Given a finite group G of order > 1 and an integer $n \geq 2$, there exist infinitely many non-homeomorphic connected fixed-point-free graphs X of chromatic number n whose automorphism group is isomorphic to G.*

Proof. Case (1): $n = 2$. Let X be as in (3.5), case (1). Every circuit of X is of even order; hence by a well-known theorem **(4**, p. 170**)**, $\chi(X) = 2$.

Case (2). $n \geq 3$. Let P_i, $i = 1,\ldots,n$, be the graph with

$$V(P_i) = \{p_1,\ldots,p_i\},\quad E(P_i) = \{[p_j,p_{j+1}],\ j = 1,\ldots,i-1\}.$$

Consider the complete n-graph $C^{(n)}$. Denote its vertices by $x_1,\ldots,x_n$. Identify the vertex x_i of $C^{(n)}$ with the vertex p_i of P_i, $i = 1,\ldots,n$. The graph C_n so obtained is prime (since it is connected, and contains vertices which do not belong to any 4-circuit of C_n), has chromatic number $\chi(C_n) = n$, and $G(C_n) = 1$.

Let X be as in case (1). Then by (2.10), $G(X \times C_n) \cong G$, by (2.9), $X \times C_n$ is fixed-point-free; and by (2.6), $\chi(X \times C_n) = n$.

(3.7) THEOREM. *Given a finite group G of order > 1 and an integer $n \geq 3$, there exist infinitely many non-homeomorphic connected fixed-point-free graphs X which are regular of degree n, and whose automorphism group is isomorphic to G.*

Proof. For $n = 3$ part of (3.7) has been proved in **(3)**. The proof given here for $n \geq 4$ is patterned after that of **(3**, Theorem 4.1**)**.

We first show that there exists an infinite sequence of cyclically connected non-isomorphic prime graphs Y_1, Y_2, ..., which are regular of degree 3, and for which $G(Y_i) = 1 (i = 1,2,\ldots)$. By **(3**, Theorem 2.3**)** there exists at least one such graph Y_1. By induction, let $Y_{i+1} = \bar{Y}_i$, $i \geq 1$. Then by (3.2), (3.3) the Y_i's have the required properties.

Let X be a fixed-point-free graph of degree n which is relatively prime to $Y_1, \ldots, Y_k$, and such that $G(X) \cong G$, where G is a given finite group of order > 1. By (2.9),

$$W_k = X \times Y_1 \times \ldots \times Y_k$$

is fixed-point-free and by (2.10), $G(W_k) \cong G$. By (2.5), W_k is regular of degree $n + 3k$. Hence (3.7) is proved if we show the following: There exist infinitely many non-isomorphic connected fixed-point-free graphs $X_j^{(n)}$ $(j = 1,2, \ldots)$, which are regular of degree $n = 3,4,5$, relatively prime to all Y_i, and such that $G(X_j^{(n)}) \cong G$ for all j.

Let $G = \{\tau\}$, and let $X_1^{(3)}$ be the graph given in (3, Theorem 4.1). $V(X_1^{(3)}) = \{x_j^\tau, j \leqslant m, \tau \in G\}$, where $m = 2h + 4$, $E(X_1^{(3)})$ as given in (3, p. 374), by quadratic forms. Define $X_1^{(4)}$, $X_1^{(5)}$ as follows:

$$\begin{aligned}
V(X_1^{(4)}) &= V(X_1^{(3)}) \cup \{y_j^\tau, j \leqslant m, \tau \in G\},\\
E(X_1^{(4)}) &= E(X_1^{(3)}) \cup \{[x_j^\tau, y_j^\tau]\ (j \leqslant m), [y_j^\tau, y_{j+1}^\tau],\ (j \leqslant m-1),\\
&\qquad [y_j^\tau, y_{m-j+1}^\tau]\ (j \leqslant h+1)\ [y_1^\tau\ y_{h+2}^\tau], [y_{h+2}^\tau, y_m^\tau],\\
&\qquad \tau \in G\};\\
V(X_1^{(5)}) &= V(X_1^{(4)}) \cup \{z_j^\tau\ \ (j \leqslant m), \tau \in G\},\\
E(X_1^{(5)}) &= E(X_1^{(4)}) \cup \{[x_j^\tau, z_j^\tau], [y_j^\tau, z_j^\tau]\ (j \leqslant m), [z_j^\tau, z_{j+1}^\tau]\\
&\qquad (j \leqslant m-1), [z_j^\tau, z_{m-j+1}^\tau]\ (j \leqslant h+1), [z_1^\tau, z_{h+2}^\tau], [z_{h+2}^\tau, z_m^\tau], \tau \in G\}.
\end{aligned}$$

It is easily checked that $X_1^{(n)}$ $(n = 4,5)$ is of degree n, and that if $\phi \in G(X_1^{(n)})$, and

$$\phi x_1^{\tau_0} = x_1^{\tau_0}$$

for some $\tau_0 \in G$, then $\phi x_j^\tau = x_j^\tau$, $\phi y_j^\tau = y_j^\tau$, $\phi z_j^\tau = z_j^\tau$, for $j \leqslant m$ and $\tau \in G$. An argument similar to that in (3) then shows that $G(X_1^{(n)}) \cong G (n = 3,4,5)$.

By induction, let $X_{j+1}^{(n)} = \bar{X}_j^{(n)}$ $(j \geqslant 1, n = 3,4,5)$. Then by (3.2), (3.3), $X_{j+1}^{(n)}$ is prime, regular of degree n, fixed-point-free, and

$$G(X_{j+1}^{(n)}) \cong G(X_j^{(n)}) \cong G, \qquad j \geqslant 1.$$

Clearly $X_j^{(n)}$ and Y_i are non-isomorphic for $i,j = 1,2, \ldots,$ and $n = 3,4,5$; hence the $X_j^{(n)}$ are relatively prime to the Y_i, and this is what we set out to prove.

(3.8) THEOREM. *Let Y be a connected graph, and let G be a finite group of order > 1. Then there exist infinitely many non-homeomorphic fixed-point-free graphs X such that* (i) $G(X) \cong G$, *and* (ii) *X is spanned by a graph $\bar{Y}$ homeomorphic to Y.*

Proof. Let $V(Y) = \{y_1, \ldots, y_r\}$. Take a spanning tree T of Y. Let e_1 be an edge of T incident with y_1. Subdivide e_1 by a new vertex z_1. Let T_1, Y_1 be the graphs obtained by this subdivision from T and Y respectively. Let

e_2 be an edge of T_1 incident with y_2. Subdivide e_2 by a new vertex z_2, obtaining graphs T_2, Y_2. Continuing this process we finally obtain a graph Y_r with

$$V(Y_r) = \{y_1,z_1,y_2,z_2,\ldots,y_r,z_r\},\ [y_i,z_i] \in E(Y_r), \qquad i \leqslant r.$$

Define H, $\bar{Y}$, Z by

$$V(H) = V(\bar{Y}) = V(Z) = V(Y_r) \cup \bigcup_{i=1}^{r} \{y_{i1},\ldots,y_{is_i}\},$$

$$E(H) = E \cup \{[y_1,z_1],\ [z_1,y_2],\ [y_2,z_2],\ldots,[z_{r-1},y_r],\ [y_r,z_r],\ [z_r,y_1]\},$$

$$E(\bar{Y}) = E \cup E(Y_r),\ E(Z) = E(H) \cup E(Y_r),$$

where

$$E = \bigcup_{i=1}^{r} \{[y_i,y_{i1}],\ [y_{i1},y_{i2}],\ldots,[y_{is_i-1},y_{is_i}],\ [y_{is_i},z_i]\},$$

$s_1,\ldots,s_r$ being positive integers to be chosen as specified below. Clearly $\bar{Y}$ spans Z and is homeomorphic to Y. H is a Hamiltonian circuit of Z, and $E(H) \cap E(\bar{Y}) \neq \square$. In Z each z_i is of degree 3. Let P_i be that path of H which joins z_i and z_{i+1}, and contains y_{i+1} (subscripts to be taken modulo r). All vertices of P_i except z_i, z_{i+1}, and possibly y_{i+1} are of degree 2 in Z. Let the s_i be so chosen that (1) $s_i > a$, where a is a given positive integer, and (2) $\alpha_0(P_{i+1}) > \alpha_0(P_i) > \alpha_1(P)$, for $i = 1,\ldots,r-1$, and all paths P of Z not containing a vertex y_{ij}. It follows from (2) that Z is prime (since no y_{ij} belongs to a 4-circuit of Z), and that $G(Z) = 1$.

Given a finite group G of order > 1, let X be a graph with the properties stated in (3.4), and let Z be the graph constructed above with $a = 2\alpha_0(X) - 2$. Then X, Y, Z satisfy the hypotheses of (2.7), and it follows that $X \times Z$ is spanned by a graph $\bar{Y}$ homeomorphic to Y. By (2.9), $X \times Z$ is fixed-point-free. X and Z are non-isomorphic, and since both graphs are prime, $G(X \times Z) \cong G(X) \cong G$.

References

1. G. A. Dirac, *The structure of k-chromatic graphs*, Fund. Math., *40* (1953), 42–55.
2. R. Frucht, *Herstellung von Graphen mit vorgegebener abstrakter Gruppe*, Compositio Math., *6* (1938), 239–250.
3. R. Frucht, *Graphs of degree 3 with given abstract group*, Canad. J. Math, *1* (1949), 365–378.
4. D. König, *Theorie der endlichen und unendlichen Graphen*, (Leipzig, 1936).
5. G. Sabidussi, *Graph multiplication*. Submitted to Mathematische Zeitschrift.
6. H. Whitney, *Congruent graphs and connectivity of graphs*, Amer. J. Math., *54* (1932), 150–168.

University of Minnesota
and
Tulane University

THE EQUIVALENCE OF QUADRATIC FORMS

G. L. WATSON

1. Introduction. The main object of this paper is to find the number of classes in a genus of indefinite quadratic forms, with integral coefficients, in $k > 4$ variables, distinguishing for even k two cases, according as improper equivalence is or is not admitted. (Two forms are in the same genus, according to the classical definition of Minkowski, if either is equivalent, for every positive integer m, to one identically congruent to the other modulo m.) Meyer **(5)** considered this problem, but obtained only a very incomplete result, included in Theorem 4 below. Otherwise little was known till recently. The results I prove could perhaps be obtained by suitable specialization of the very deep work of Eichler **(2)**; but it seems worth while to give a more elementary treatment of the case when the coefficients and variables are in the ring of ordinary integers.

The present paper may be regarded as a sequel to **(3)**, which gives the result for $k = 3$. It is, however, independent of **(3)** in so far as the results for $k > 4$ are concerned. It turns out that the formula giving the exact value of the class-number (in either sense) for an indefinite form with $k > 3$ gives a lower bound for that of any form with $k > 3$. The forms considered are not therefore assumed to be indefinite unless so stated; nor (since the proofs are partly by induction on k) to have $k > 4$.

2. Notation. Small letters denote rational integers unless otherwise stated, p being a prime and $(n|p)$ $(p \neq 2)$ the Legendre symbol. (m, n) denotes as usual the greatest common divisor of m, n. The set of all square-free integers $(v, v_1, \dots)$ constitutes, with the operation

$$v_1 \cdot v_2 = v_1 v_2 (v_1, v_2)^{-2}, \tag{2.1}$$

a group, denoted by Γ. Any subset of Γ closed under this operation is a subgroup; so in particular is Γ_d, the subset with $(v, d) = 1$.

Latin capitals denote square matrices, of rank k unless otherwise indicated, with rational elements, I being the identity matrix. By the denominator of a matrix is meant the least common multiple of the denominators of its elements, and the determinant is denoted by modulus signs. The notation $[m_1, \dots, m_k]$ is used for a diagonal matrix; and similarly for a matrix made up of diagonal blocks. Transposition is indicated by an accent. Column vectors, or $k \times 1$ matrices, are written $\mathbf{x} = \{x_1, \dots, x_k\}$ and have integral elements unless otherwise stated.

Received September 6, 1956; in revised form April 24, 1957.

Congruences, vector or scalar, in which either side is fractional, but with denominator prime to the modulus, are to be interpreted in the usual way. $m|n$, $m \nmid n$, $p^\alpha||n$ denote respectively that m divides n, m does not divide n, p^α divides n but $p^{\alpha+1}$ does not.

3. The matrix and discriminant of a form. These are defined (see, for example, Brandt, **1**) without putting in the Gaussian binomial coefficients. That is, $a_{ij} = a_{ji}$ is the coefficient of $x_i x_j$ in $f(\mathbf{x}) = f(x_1, \ldots, x_k)$, and with the form f we associate the matrix

$$A = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right)$$

with elements $2a_{ii}$ and a_{ij} $(i \neq j)$. This gives $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}'A\mathbf{x}$ in place of the "classical" $\mathbf{x}'A\mathbf{x}$. Since f is assumed to have integral coefficients, A has integral elements, those on its diagonal being even. It is thus congruent (mod 2) to a skew matrix, which for odd k is singular. The discriminant of f, defined by

$$d = d(f) = \begin{cases} (-1)^{\frac{1}{2}k}|A| & \text{for } k \text{ even} \\ \frac{1}{2}(-1)^{\frac{1}{2}(k-1)}|A| & \text{for } k \text{ odd} \end{cases}$$

is therefore always integral; and we assume always that f is not degenerate, that is, that $d \neq 0$.

If p^β is any power of a prime p not dividing d, then by a suitable integral unimodular transformation we may suppose **(1)** that, for odd k,

$$3.1 \qquad f(\mathbf{x}) \equiv x_1x_2 + \ldots + x_{k-2}x_{k-1} + dx_k^2 \pmod{p^\beta},$$

or for even k,

$$3.2 \qquad f(\mathbf{x}) \equiv x_1x_2 + \ldots + x_{k-2}x_{k-2} + \phi \pmod{p^\beta},$$

where ϕ is any binary form, with discriminant d, in x_{k-1}, x_k. Similarly we may suppose **(6)** for any odd p^β, whether or not p divides d, that

$$3.3 \qquad f(\mathbf{x}) \equiv \sum_{i=1}^{k} p^{\lambda_i} a_i x_i^2 \pmod{p^\beta}, \qquad p \nmid a_1 \ldots a_k,$$

where the exponents λ_i may be supposed arranged in ascending order. For $p = 2$ we must replace 3.3 by **(6**, 35, Lemma 3**)**

$$3.4 \qquad f(\mathbf{x}) \equiv \sum_{\rho=1}^{\nu} 2^{\mu_\rho}\phi_\rho(x_{2\rho-1}, x_{2\rho}) + \sum_{i=2\nu+1}^{k} 2^{\lambda_i} a_i x_i^2 \pmod{2^\beta}.$$

Here $0 \leqslant \nu \leqslant \frac{1}{2}k$, the a_i are odd, and the binary forms ϕ_ρ have odd discriminants d_ρ. The properties of such a form depend on the residue (1 or -3) of d_ρ (mod 8); but we shall see that this distinction is irrelevant for our purpose.

From 3.1, 3.2 we see that the arithmetical properties of f to a modulus prime to d are trivial; they are given uniquely when k and d are known. The properties of f to any modulus may thus be studied by means of 3.3, 3.4, with p ranging over the divisors of d; and it is convenient to replace this system of

congruences by a single one, with a power of d as modulus; we shall see that the fourth power of d is high enough. Combining the results 3.3, 3.4 (for $p|d$) we see that we may suppose

$$3.5 \qquad f(\mathbf{x}) \equiv \sum_{\rho=1}^{\nu} \{q_{2\rho-1}n_{2\rho-1}(x_{2\rho-1} - \tfrac{1}{2}x_{2\rho})^2 + \tfrac{1}{4}q_{2\rho}n_{2\rho}x_{2\rho}^2\} + \sum_{i=2\nu+1}^{k} q_i n_i x_i^2 \pmod{d^4}.$$

Here the n_i are products of primes dividing d, while the q_i may without loss of generality be taken to be in Γ_{2d}; ν is as in 3.4 if d is even, 0 otherwise; and for $\rho = 1, \ldots, \nu$ we must have

$$2^{\mu_\rho} || n_{2\rho} \equiv -q_{2\rho-1}q_{2\rho}n_{2\rho-1} \pmod{2^{\mu_\rho+2}},$$

the expression in $\{\}$ being a binary form with odd discriminant, multiplied by

$$2^{\mu_\rho}.$$

Alternatively, we might obtain 3.5 by the same elementary method (essentially completing the square) which gives 3.3, 3.4.

We see from 3.5 that

$$3.6 \qquad d \equiv (-4)^{[\frac{1}{2}k]}4^{-\nu}(q_1 \ldots q_k)(n_k \ldots n_k) \pmod{d^4},$$

whence $4^{[\frac{1}{2}k]-\nu}n_1 \ldots n_k$ is a divisor of d; so since the q_i are prime to d we must have

$$3.7 \qquad q_1 \ldots q_k \equiv \pm 1 \pmod{d^3}.$$

4. The groups and automorphs of a form. We define

$$4.1 \qquad U(\mathbf{t}) = U(\mathbf{t}, f) = U(\mathbf{t}, A) = I - \mathbf{t}\mathbf{t}'A/f(\mathbf{t}),$$

for $\mathbf{t}$ with $f(\mathbf{t}) \neq 0$. This matrix (which is $-U(\mathbf{t})$ in the notation of (3)) is well known, and may be immediately verified, to be an automorph of f, or of A. That is, we have identically $f(U(\mathbf{t})\mathbf{x}) = f(\mathbf{x})$. The first two of the following formulae are immediate consequences of 4.1, and as they show that $U(\mathbf{t})$ has linearly independent characteristic vectors with characteristic roots -1, $1, \ldots, 1$, the other two follow:

$$4.2 \qquad U(\mathbf{t})\mathbf{t} = -\mathbf{t}; \; U(\mathbf{t})\mathbf{x} = \mathbf{x}, \text{ if } \mathbf{t}'A\mathbf{x} = 0; \quad |U(\mathbf{t})| = -1; \; U^2(\mathbf{t}) = I.$$

Since 4.1 gives $U(n\mathbf{t}) = U(\mathbf{t})$ for $nf(\mathbf{t}) \neq 0$, we may allow fractional $\mathbf{t}$, and then we have for all non-singular R

$$4.3 \qquad U(R^{-1}\mathbf{t}, R'AR) = R^{-1}U(\mathbf{t}, A)R;$$

that is, any linear transformation takes U's into U's.

On the other hand, if we take $\mathbf{t}$ to be integral and primitive, that is, assume that the greatest common divisor of $t_1, \ldots, t_k$ (all integers) is 1, then some linear combination of the rows $t_i\mathbf{t}'A$ of the matrix $\mathbf{t}\mathbf{t}'A = (t_it_j)A$ is $\mathbf{t}'A$.

Hence if n is the greatest common divisor of $f(\mathbf{t})$ and the k elements of $\mathbf{t}'A$, then $n^{-1}f(\mathbf{t})$ is the denominator of $U(\mathbf{t})$. We are interested in $U(\mathbf{t})$ with denominator prime to d. The residue modulo d^4 of the denominator of $U(\mathbf{t})$ will be considered first.

We consider the n, $\mathbf{t}$, q satisfying

4.41 $\quad n|f(\mathbf{t}),$

4.42 $\quad n|\mathbf{t}'A,$

4.43 $\quad n|d,$

4.44 $\quad q \in \Gamma_d,$

4.45 $\quad f(\mathbf{t}) \equiv qn \pmod{d^4},$

4.46 $\quad qnf > 0$ if f is definite.

4.46 means that qn has the sign of f if f is a definite form; otherwise qn may be either positive or negative. These conditions 4.4 will be studied further in §5; meanwhile we define certain groups.

Definition of $\Gamma(f)$, $\Gamma^+(f)$. $\Gamma(f)$ is the subgroup of Γ_d generated by the set of q for which, for suitable integral $n = n(q)$, $\mathbf{t} = \mathbf{t}(q)$, conditions 4.4 can be satisfied. $\Gamma^+(f)$ is the subgroup (of index 1 or 2) of $\Gamma(f)$ generated by the products in Γ of pairs of such q.

There is a connection, which we shall investigate in §5, between conditions 4.4 and

4.51 $\quad p^\delta|f(\mathbf{t}),$

4.52 $\quad p^\delta|\mathbf{t}'A,$

4.53 $\quad p^\delta|d,$

4.54 $\quad f(\mathbf{t}) \equiv b \pmod{p^{\delta+2}}.$

Definition of $\Gamma(p, f)$. $\Gamma(p, f)$ is the subgroup of Γ generated by the set of v given by

4.6 $\quad b_1b_2 = u^2v,\ u$ integral, $v \in \Gamma,$

where b_1, b_2 range independently over the set of b for which, for given p and suitable $\delta = \delta(b) > 0$, and integral $\mathbf{t} = \mathbf{t}(p)$, 4.5 can be satisfied.

When $p \nmid d$, 4.53 gives $\delta = 0$, while 4.54 is soluble (unless $k = 1$) for every b not divisible by p, as may be seen from 3.1 or 3.2; hence

4.7 $\quad \Gamma(p, f) = \Gamma_p$ if $p \nmid d.$

Here Γ_p (see §2) is the subgroup of Γ defined by $(v, p) = 1$.

All these groups are clearly unaltered if f is replaced

(i) by any equivalent form, or

(ii) by any form congruent to $f \pmod{d^4}$, and with the same signature and discriminant.

(To deduce (ii), note that 4.53 gives $p^{\delta+2}|d^4$ if $p|d$, and use 4.7 if $p \nmid d$.) It follows from the Minkowski definition that the groups are all invariants of the genus of f. We can now state our main result:

THEOREM 1. *Let f be a non-degenerate quadratic form, with integral coefficients, in at least three variables. Then* (i) *the number of classes in the genus of f is not less than the order of the factor group $\Gamma_{d(f)}/\Gamma(f)$ or $\Gamma_{d(f)}/\Gamma^+(f)$, according as improper equivalence is or is not admitted*;

(ii) $\Gamma(f) = \Gamma^+(f)$ *is a necessary condition for f to be improperly equivalent to itself*;

(iii) *if f is indefinite, then there is equality in* (i) *and the necessary condition in* (ii) *is also sufficient.*

It is clear that the value, or lower bound, given by this theorem for the class-number, in either sense, is always a power of 2

5. Relations between the groups. We show first that, for primitive $\mathbf{t}$, 4.41 and 4.42 imply 4.43; whence, if $p \nmid \mathbf{t}$, 4.51 and 4.52 imply 4.53. We may suppose, by an integral unimodular transformation, that $\mathbf{t} = \{1,0,\ldots,0\}$. Then 4.41, 4.42 reduce to

$$n|a_{11},\ n|\{2a_{11}, a_{12}, \ldots, a_{1k}\}.$$

And the substitution $\mathbf{x} \to U(\mathbf{t})\mathbf{x}$ reduces to

$$x_1 \to -x_1 - a_{11}^{-1}(a_{12}x_2 + \ldots).$$

The leading element $2a_{11}$ of A is divisible by $2n$, and its first row and column by n, so $|A|$ is divisible by $(2n, n^2) = n$ or $2n$ according as n is odd or even, giving $n|d = \pm|A|$ or $\pm\frac{1}{2}|A|$.

It follows now that when 4.4 holds $U(\mathbf{t})$ has denominator $n^{-1}f(\mathbf{t}) \equiv q \pmod{d^2}$. For if not, then 4.41, 4.42 would hold also with np, $p^{-1}q$ for n, q, p a prime not dividing d, whence $np \nmid d$. It also follows that the possibilities for q are the same whether or not $\mathbf{t}$ in 4.4 is restricted to be primitive. Similarly, it does not matter whether or not we allow $\mathbf{t}$ in 4.5 to be divisible by p; the possibilities for b are the same in either case, up to a square factor, which in view of 4.6 does not matter.

To reconcile the definition of $\Gamma(p, f)$ with that given, for $k = 3$, in (3), we show that, for odd k, 4.6 may be replaced, in the definition of $\Gamma(p, f)$, by

$$5.1 \qquad (-1)^{\frac{1}{2}(k-1)}db = u^2v,\ u \text{ integral},\ v \in \Gamma.$$

For $p \nmid d$ this is clear from 4.7. If $p \mid d$, we note that the numbers $q_i n_i$ of 3.5 are admissible values of b; the corresponding $\mathbf{t}$ are the vectors making all but one of the squares in 3.5 vanish. It is clear that in 4.6 we may allow b_2 alone to vary, and replace b_1 by a fixed product of an odd number of b. Using the b just found, 3.6 gives the desired result.

Similarly we show that

$$5.2 \qquad \Gamma(f) = \Gamma^+(f), \qquad k \text{ odd}.$$

Since every f is trivially equivalent to itself by $\mathbf{x} \to -\mathbf{x}$, which is an improper equivalence for odd k, 5.2 shows that the assertions of Theorem 1 simplify

as they should for odd k, the distinction between proper and improper equivalence disappearing.

To prove 5.2, note that the n_t, q_t of 3.5 satisfy 4.4, with the same $\mathbf{t}$ as used in connection with 5.1. Their group product, together with $\Gamma^+(f)$, obviously generates $\Gamma(f)$. Hence 5.2 follows if we show that this group product, which by 3.7 is either a quadratic residue modulo d^3 or the negative of such a residue, is in $\Gamma^+(f)$. Now $\Gamma^+(f)$ contains -1, since we may put $-n$, $-q$ for n, q in 4.4; it also contains all quadratic residues modulo d^3, as we see by keeping n fixed in 4.4 and putting $m\mathbf{t}$ for $\mathbf{t}$, m prime to d, and $q' \equiv m^2q$ modulo d^3 for q. 5.2 follows.

The relation between $\Gamma^+(f)$ and the groups $\Gamma(p, f)$ is given by

LEMMA 1. *q is in $\Gamma^+(f)$ if and only if, for suitable $w = w(q)$ in Γ,*

$$q \in \Gamma_d, \tag{5.31}$$

$$w|d, \tag{5.32}$$

$$wq \in \bigcap_{p|d}{}' \ \Gamma(p,f). \tag{5.33}$$

where the accent denotes the exclusion of negative values of wq in case f is definite.

Proof. We note first that the set of q for which 5.3 can be satisfied is a group, say $\Gamma_+(f)$; for if w_1, q_1 and w_2, q_2 satisfy 5.3 then so do $w_1 \cdot w_2$ and $q_1 \cdot q_2$.

Now note that 4.4 implies 4.5 (with $b = qn$) for every p dividing d. For $p|d$ and $p^\delta|d$ together imply $p^{\delta+3}|d^4$. Hence, writing $n = wc^2$, $w|d$, in 4.4, we see that the "only if" of the lemma, that is, $\Gamma^+(f) \subseteq \Gamma_+(f)$, follows from the definitions of $\Gamma^+(f)$ and $\Gamma(p, f)$. (Note that the product of evenly many qn, or qw, of the same sign is always positive.)

Now to prove the "if," that is, $\Gamma_+(f) \subseteq \Gamma^+(f)$, we consider integers v in Γ with the property that, for each $p|d$ and suitable u_p, $u_p{}^2v$ is an admissible value of b in 4.5, while vf is positive if f is definite. It is clear that products of pairs of such v generate the group on the right of 5.33, while the corresponding products of pairs of values of $\pm(v, d)^{-1}v$ generate $\Gamma_+(f)$. It suffices therefore to show that to each such v there is a u such that 4.4 can be satisfied with $qn = u^2v$. Now the condition that 4.5 can be satisfied with $b = u_p{}^2v$ is obviously satisfied, if at all, with u_p a power of p. So we suppose u_p is a power of p, and write $u = \Pi_p u_p$; clearly u^2v is an admissible value of b in 4.5 for every p dividing d. This is still true, by elementary properties of quadratic residues, if the exponent $\delta + 3$ in 4.54 is replaced by β such that $p^\beta || d^4$. Comparing 4.5, as thus modified, with 4.4 we see easily that, with $q = \pm\ (v, d)^{-1}v$, $qn = n^2v$, we can satisfy 4.41 to 4.45. And as 4.46 is satisfied (if applicable) by our choice of the sign of v, the proof is complete.

From 5.1 and Lemma 1 it follows that the group $\gamma(f)$ of **(3)** coincides with $\Gamma(f)$ and with $\Gamma^+(f)$ when $k = 3$ and f is indefinite.

Theorem 1 is true for imprimitive forms, and we shall later need to be free to exclude such forms or not, as convenient. So we prove that *the factor groups*

of Theorem 1 *are unaltered up to isomorphism if* f *is replaced by* mf, $m \neq 0$. We do this by showing that on so doing each of $\Gamma(f)$, $\Gamma^+(f)$, $\Gamma_{d(f)}$ is replaced by a subgroup of itself of index 2^σ, σ being the number of primes dividing m but not d. As far as $\Gamma_{d(f)}$ is concerned this is clear. For the other two groups it suffices to show that on the one hand q can satisfy 4.4 and have all or any of these p as divisors (which is trivial), while on the other hand if 4.4 holds with $(q, m) = 1$ it also holds with mn, q, mf for n, q, f. This last assertion depends on modifying the choice of $\mathbf{t}$ so as to satisfy 4.45 to a higher modulus; this is done as in the proof of Lemma 2 below, and is straightforward.

A similar argument gives

$$\Gamma(p, mf) = \Gamma(p, f) \text{ for } m \neq 0. \tag{5.6}$$

6. Construction of automorphs. To obtain an upper bound for the class-number, we need to construct automorphs, of the special type 4.1, with convenient properties; in particular, with equality in 4.45. We prove:

LEMMA 2. *Suppose that* f *is indefinite,* $k > 4$, 4.4 *holds, and also* $f(\mathbf{t}) \equiv qn \pmod{q^2}$. *Then there exists* $\mathbf{z}$ *satisfying*

$$\mathbf{z} \equiv \mathbf{t} \pmod{dq^2}, f(\mathbf{z}) = qn. \tag{6.1}$$

Proof. With the present hypotheses, and $d \neq 0$, it suffices, by the result proved in **(7)**, to show that a solution of 6.1 is not excluded by congruence considerations. In other words, we need only show that

$$\mathbf{z} \equiv \mathbf{t} \pmod{dq^2}, f(\mathbf{z}) \equiv qn \pmod{m} \tag{6.2}$$

can be satisfied for any prescribed $m \neq 0$, and suitable $\mathbf{z}$.

Suppose first $(m, dq) = 1$; then since $k > 2$ there is at least one product term in 3.1 or 3.2 (for any prime power factor of m) and so 6.2 is trivial. Next suppose $m = d^s$, $s \geq 5$. We can find $r \equiv 1 \pmod{d^3}$ so that $qn \equiv rf(\mathbf{t}) \pmod{d^s}$. Then we can solve $h^2 \equiv r \pmod{d^s}$; and it suffices to put $\mathbf{z} = h\mathbf{t}$. We may therefore suppose $m = q^s$, $s \geq 3$. Proceeding by induction on s, suppose $\mathbf{z} = \mathbf{x}$ satisfies 6.2 with $m = q^{s-1}$. Put $\mathbf{z} = \mathbf{x} + q^{s-1}\mathbf{y}$; then 6.21 holds, and 6.22 with $m = q^s$ reduces to

$$f(\mathbf{z}) \equiv f(\mathbf{x}) + q^{s-1}\mathbf{x}'A\mathbf{y} \equiv f(\mathbf{x}) + q^{s-1}\mathbf{t}'A\mathbf{y} \equiv qn \pmod{q^s}$$

This reduces to a linear congruence of the type $\mathbf{t}'A\mathbf{y} \equiv l \pmod{q}$, which is soluble for $\mathbf{y}$ unless some p dividing q, hence not dividing d, by 4.44, divides $\mathbf{t}'A$, and also, by hypothesis, $f(\mathbf{t})$. If so, then with $n = p$ 4.41 and 4.42 hold and 4.43 fails, which as shown at the beginning of §5 is impossible.

We deduce the

COROLLARY. *Suppose* f *is indefinite,* $k > 4$, *and* 4.4 *holds. Then there exists* $\mathbf{z}$ *with* $f(\mathbf{z}) = qn$ *such that* $U(\mathbf{z})$ *has denominator* q *and satisfies*

$$qnU(\mathbf{z}) = (p^2\xi_2, p\xi_2, \ldots, p\xi_{k-1}, \xi_k) \tag{6.3}$$

with integral ξ_i, for every p dividing q for which

6.4 $$p^2 | a_{11},\ p | a_{1j}, \qquad 1 < j < k.$$

Proof. We apply the lemma with a suitable $\mathbf{t}$. If p divides q but does not satisfy 6.4, any solution of $f(\mathbf{t}) \equiv qn \pmod{p^2}$ will do, and some solution clearly exists, by 3.1 or 3.2. If p divides q and satisfies 6.4, we require a solution of $f(\mathbf{t}) \equiv qn \pmod{p^2}$ which also satisfies

6.5 $$\mathbf{t} \equiv \{1, 0, \ldots, 0, q\theta\} \pmod{p^2},$$

for some θ. The congruence $f(\mathbf{t}) \equiv qn \pmod{p^2}$ reduces, by 6.4, 6.5, to $a_{1k}q\theta \equiv qn \pmod{p^2}$, which is soluble since $p \nmid a_{1k}$. For $p | a_{1k}$ would with 6.4 give $p | d$, $(q, d) > 1$, contradicting 4.4.

Now 6.3 follows from 4.1, 6.4, 6.5 by a simple calculation.

7. Rational transformations. Denote by R a matrix, with determinant ± 1 and denominator prime to $d(f)$, which takes f into $f^R = f(R\mathbf{x})$ with integral coefficients. Impose for the moment, in case $d(f)$ is odd, the additional restriction that the denominator of R be odd. Then it is well known that every form in the genus of f is expressible as f^R, and conversely. (This is equivalent to saying that the Eisenstein-Smith definition of the genus by rational transformations is equivalent to that of Minkowski, which we have used.) The additional restriction on the denominator of R is easily removed, as we shall see.

Among the matrices R are included all automorphs S of f whose denominators are prime to $d(f)$, and also all products SR, since $f^{SR} = f^R$. For given R, we shall construct S so that if possible SR is simpler, in a sense to be defined, than R. The construction requires f to be indefinite, and $k \geqslant 4$.

It is not difficult to express R as

7.1 $$R = T[r_1 s_1^{-1}, \ldots, r_k s_k^{-1}]X, \quad |T| = 1, \quad |X| = |R|,$$

where the matrices T, X are integral, and the positive integers r_i, s_i satisfy $(r_i, s_i) = 1$ and

7.2 $$1 = r_1 | r_2 | \ldots | r_k, \quad 1 = s_k | s_{k-1} | \ldots | s_1.$$

The proof that this is possible is similar to the proof that 7.3, below, is possible, and so we omit it; but we note that the r_i, s_i depend only on R and not on T, X. For firstly, $r_1 s_1^{-1}$ is the largest positive rational fraction such that $r_1^{-1} s_1 R$ has integral elements. Next, $r_1 r_2 s_1^{-1} s_2^{-1}$ is the largest fraction such that all the 2×2 submatrices of R have determinants which are integral multiples of $r_1 r_2 / s_1 s_2$; and so on.

Similarly, we can for any p express R as

7.3 $$R = M[p^{\theta_1}, \ldots, p^{\theta_k}]N, \quad |M| = 1, \quad |N| = |R|, \qquad \theta_1 \leqslant \theta_2 \leqslant \ldots \leqslant \theta_k,$$

where M, N have denominators prime to p, and the integers $\theta_i = \theta_i(R, p)$ depend only on i, R, p and not on M, N. The sum of these integers is obviously

zero, and they all vanish if p does not divide the denominator of R. To prove that R can be expressed in the form 7.3, suppose for the moment that M satisfying 7.32 has been suitably chosen. Then $\theta_1, \ldots, \theta_k$ and N may be chosen so that 7.31 holds, by simply taking out from each row vector of $M^{-1}R$ the highest possible power of p so as to leave a vector with denominator prime to p. If now 7.33 fails, then p must divide $|N|$, and we see that a higher power of p can be taken out from one of the rows after a suitable row operation on $M^{-1}R$, equivalent to a suitable modification of the choice of M.

Definition. We define $q(R)$ to be the product of all primes p for which the sum of the positive ones among the numbers $\theta_i(R, p)$ is odd.

LEMMA 3. *The integers $\theta_i = \theta_i(R, p)$ of 7.33 satisfy*

$$\theta_i = -\theta_{k+1-i}, \qquad 1 \leqslant i \leqslant k \tag{7.4}$$

and the r_i, s_i of 7.1 satisfy $r_i = s_{k+1-i}$, whence $r_i = 1$ for $i \leqslant \frac{1}{2}k$, and 7.1 may be rewritten

$$R = T[s_1^{-1}, \ldots s_l^{-1}, 1, s_l, \ldots, s_1]X, \tag{7.5}$$

where T, X are integral, $l = [\frac{1}{2}k]$, and the 1 is to be omitted for even k.

Proof. We use for the first time the hypothesis that f^R is integral. We suppose $p \nmid d$; otherwise the θ_i are all zero and 7.4 is trivial. It suffices to prove 7.4 since the remaining assertions follow easily. Suppose 7.4 false and let h ($\leqslant \frac{1}{2}k$) be the least i for which it fails. Suppose also $\theta_h + \theta_{k+1-h} < 0$; for otherwise we may replace f, R, θ_i by f^R, R^{-1}, $-\theta_{k+1-i}$.

Suppose further that $T = I$ in 7.1; for if not we may replace f, R by f^T, $T^{-1}R$. Then it easily follows that we can take $M = I$ in 7.3. It is easily seen that

$$f^{RN^{-1}}$$

is also integral, so we may suppose

$$R = [p^{\theta_1}, \ldots, p^{\theta_k}].$$

The coefficient of $x_i x_j$ in f^R is now $p^{\theta_i+\theta_j} a_{ij}$; so we must have

$$p|a_{ij} \text{ if } \theta_i + \theta_j < 0.$$

By 7.34 and our hypotheses regarding the θ_i, this gives

$$p|a_{ij} \text{ for } i \leqslant h,\ j \leqslant k + 1 - h.$$

This shows that the matrix A_0 derived from A by replacing by zero every element $2a_{ii}$ or a_{ij} ($i \neq j$) with $p|a_{ij}$ has rank $< k$. For the submatrix of A_0 consisting of its first h rows has rank $< h$.

Thus $|A_0| = 0$. If $p \neq 2$, this gives the contradiction $|A| \equiv |A_0| \equiv 0 \pmod{p}$, $p|d$. If $p = 2$, we may have $|A| = \pm 2d$, so we need to prove $|A| \equiv |A_0| \pmod 4$. If we consider the terms in the expansion of $|A|$ that vanish in that of $|A_0|$, we see that they are all even, and either occur in pairs

(because of the symmetry of A) or contain factors $2a_{ii}$ or $a_{ij}a_{ji} = a_{ij}^2 \equiv 0 \pmod 4$. This completes the proof.

COROLLARY. *We may assume* $T = I$ *in* 7.5 *and simultaneously that* 6.4 *holds for every* p *dividing the denominator of* R.

Proof. We have seen in the proof of the lemma that we may assume $T = I$, and that then 6.41 holds since by hypothesis $\theta_1 < 0$. Further, we have 6.42 for every j for which $\theta_1 + \theta_j < 0$. This is true by 7.34 and 7.5 unless $j \geqslant \frac{1}{2}(k+2)$.

We consider for simplicity the case in which $a_{1,k-1}$ and $a_{1,k}$ are the only a_{1j} not divisible by p; in this case, $\theta_1 = \theta_2 = -\theta_{k-1} = -\theta_k$. We can transform f so that 6.4 holds by a suitable matrix V^{-1}, where

$$V = \begin{pmatrix} I_{k-2} & 0 \\ 0 & W \end{pmatrix},$$

W being a 2×2 matrix. But then we have to put $T = V$ instead of $T = I$ in 7.5. We thus have $R = VDX$, where D denotes the diagonal matrix in 7.5. We may write instead $R = D(D^{-1}VD)X$ if $D^{-1}VD$ is integral. Now

$$D^{-1}VD = \begin{pmatrix} I_{k-2} & 0 \\ 0 & W_1 \end{pmatrix},$$

where W_1 is derived from W by multiplying the second row, and dividing the second column, by s_1/s_2. s_1/s_2 is an integer by 7.22, divisible exactly by

$$p^{-\theta_1+\theta_2} = p^{\theta_{k-1}-\theta_k} = p^0.$$

Hence W_1 can be integral, and $W \equiv I \pmod m$ for any assigned m prime to p, without restricting W in any way modulo p. The result follows.

When R is an automorph $U(\mathbf{z})$, the numbers s_i are easily found. By 4.3, with a suitable integral unimodular matrix in place of R, we may suppose $\mathbf{z} = \{1,0,\ldots,0\}$. The positive and zero θ_i are determined by 7.4 when the negative ones are known, and the latter are clearly the same for $U(\mathbf{z})$ as for $I - U(\mathbf{z}) = \mathbf{z}\mathbf{z}'A/f(\mathbf{z})$, which has only one non-zero row. Thus θ_i is zero for $1 < i < k$, $s_i = 1$ for $i \neq 1$, s_1 is the denominator of $U(\mathbf{t})$, and $p^{-\theta_1} || s_1$, as is easily seen.

Now we consider the effect on the θ_i of replacing R by $U(\mathbf{z})R$, with suitable $\mathbf{z}$. It is convenient to write

$$\theta_i = \theta_i(R,p), \theta_i' = \theta_i(U(\mathbf{z}),p), \theta_i'' = \theta_i(U(\mathbf{z})R,p),$$

and desirable to choose $\mathbf{z}$ so that

$$q(U(\mathbf{z})R) = q(U(\mathbf{z})).q(R). \tag{7.6}$$

We prove:

LEMMA 4. (i) *For suitable* $\mathbf{z}$, $U(\mathbf{z})$ *has denominator prime to* d, 7.6 *holds, and for each* p *dividing the denominator of* R *we may as we choose have either*

(a) $\theta_i' = 0, \theta_i'' = \theta_i \quad (1 \leqslant i \leqslant k)$ *or*

(b) $(\theta_1', \ldots, \theta_k') = (-1,0,\ldots,0,1)$ *and* $(\theta_1'', \ldots, \theta_k'')$ *is a permutation of* $(\theta_1 + 1, \theta_2, \ldots, \theta_{k-1}, \theta_k - 1)$.

(ii) *If* f *is indefinite and* $k > 4$ *we may further have any positive* q *for which* 4.4 *can be satisfied as the denominator of* $U(\mathbf{z})$, *provided only that if* p *divides the denominator of* R *then* $p|q$ *if and only if alternative* (b) *is chosen in* (i).

Proof. It is convenient to assume throughout that f is indefinite and $k > 4$, and use the Corollary to Lemma 2. In other cases, when only (i) is to be proved, a suitable congruence condition may replace the Diophantine equation 6.12.

Suitable congruence conditions modulo d^4, and modulo p for every p for which we wish to satisfy (a), will clearly give $U(\mathbf{z})$ with denominator prime to d and to every such p, so that for each such p all the θ_i' vanish. Then $\theta_i'' = \theta_i$ for each such p, as we see on premultiplying 7.3 by $U(\mathbf{z})$.

For the p for which we have to satisfy (b), we use the corollaries to Lemmas 2, 3. Multiplying 6.3 and 7.31, with $T = I$, the result easily follows.

For the p which divide the denominator of $U(\mathbf{z})$ but not that of R, we have all the θ_i zero, and $\theta_i'' = \theta_i'$ is proved just like (a). Thus for every p to be considered we see that the sum of the positive θ_i'' is congruent modulo 2 to the sum of the positive θ_i and θ_i'. Plainly this gives 7.6.

Assertion (ii) is now trivial on choosing q suitably in the corollary to Lemma 2.

8. Upper bound for the class-number. We prove:

THEOREM 2. *Every form* f^R, *with* R *satisfying the conditions of the last section, is in the genus of* f.

If f *is indefinite and* $k > 4$, *then the class of* f^R, *in the wide sense, depends only on the coset of* $\Gamma(f)$, *in* Γ_d, *to which* $q(R)$ *belongs; thus in particular* f^R *is equivalent to* f *if* $q(R)$ *is in* $\Gamma(f)$.

If $|R| = 1$ *similar results, but with* $\Gamma^+(f)$ *for* $\Gamma(f)$, *hold for proper equivalence.*

Proof. The first assertion is classical for R with odd denominator. If R has an even denominator (which is possible only if d is odd) then the following argument gives $f^R = f^V$ for some V with denominator odd and prime to d.

Now let f be indefinite, and $k > 4$. Note that, since the square of every element of Γ is 1, q_1 and q_2 are in the same coset of $\Gamma(f)$ in Γ_d if and only if $q_1.q_2$ is in $\Gamma(f)$.

Denote by Q a matrix satisfying the conditions imposed in §7 on R, and in addition

8.1 $$s_1(Q) = q \in \Gamma_d,\ s_i(Q) = 1 \text{ if } i \neq 1.$$

8.1 is equivalent, by 7.4, 7.5, to

8.2 $$\theta_1(Q,p), \ldots, \theta_k(Q,p) = -1,0,\ldots,0,1,$$

for each p dividing the denominator of Q. We have seen in the proof of Lemma 4 that every $U(\mathbf{z})$ with denominator q in Γ_d is a Q with $q(U(\mathbf{z})) = q$.

Now apply Lemma 4 repeatedly. At each step the sum of the positive θ_i may be made to decrease for any p for which it exceeds 1; and since we always take $U(\mathbf{z})$ to be a Q, the sum in question does not exceed 1 for any prime factor newly introduced into the denominator. Thus after sufficiently many steps we see that we have $SR = Q$, for some $S = \ldots U(\mathbf{z}_2)U(\mathbf{z}_1)$ which is an automorph of f. We have by 7.6

$$q(Q) = q(SR) = \ldots q_2 \cdot q_1 \cdot q(R),$$

where $q_1 = q(U(\mathbf{z}_1)), \ldots$. The numbers $q_1, q_2, \ldots$ may after a certain stage, when we have already cancelled out all unwanted factors from the denominator of R, be arbitrary positive numbers that are admissible values of q in 4.4. Their product in Γ may thus, by the definition of $\Gamma(f)$, be any element of $\Gamma(f)$. We thus have $SR = Q$ with any $q(Q)$ in the coset of $\Gamma(f)$, in Γ_d, to which $q(R)$ belongs.

Now if $q(R)$ is in $\Gamma(f)$ we take $q(Q) = 1$, which by 8.1 makes Q integral, so that $f^R = f^{SR} = f^Q$ is equivalent to f. This proves the third assertion.

To prove the second assertion, take any two forms

$$f^{R_1}, f^{R_2}$$

which have $q(R^1) \cdot q(R^2)$ in $\Gamma(f)$, which are to be proved equivalent. Express them, by the foregoing construction, as

$$f^{Q_1}, f^{Q_2}$$

where $q(Q_1) \cdot q(R_1)$ and $q(Q_2) \cdot q(R_2)$ are in $\Gamma(f)$, whence $q(Q_1) \cdot q(Q_2)$ is in $\Gamma(f)$. We prove the second assertion by applying the third with f^{Q_1}, $Q_1^{-1}Q_2$ for f, R. Since f^{Q_1} is in the genus of f, it has the same groups as f, and we need only show that $q(Q_1^{-1}Q_2)$ is in $\Gamma(f)$.

It is easily seen that we may take $q(Q_2)$ to be prime to $q(Q_1)$; for $q(Q_2)$ can be any positive integer in the same coset of $\Gamma(f)$ in Γ_d as $q(R_2)$. 8.1 with 7.5 shows that Q_1^{-1} is also a Q, with denominator $q(Q_1^{-1}) = q(Q_1)$ prime to $q(Q_2)$. Hence as in the proof of Lemma 4 we see that

$$q(Q_1^{-1} Q_2) = q(Q_1^{-1}) \cdot q(Q_2) = q(Q_1) \cdot q(Q_2),$$

which is in $\Gamma(f)$, as was to be proved.

To prove the result for proper equivalence we proceed in the same way. But since the matrices $U(\mathbf{z})$ have determinant -1 by 4.23, we must premultiply by evenly many of them; the corresponding products of evenly many q given 4.4 generate $\Gamma^+(f)$.

If $k > 3$, transform f so that 3.1 or 3.2 holds, with $\beta = 2$, for each $p|q$, for suitable q in Γ_d. Then

$$Q = [q, q^{-1}, 1, \ldots, 1]$$

takes f into f^Q in the genus of f, and in a class determined by the coset of $\Gamma(f)$ or $\Gamma^+(f)$ in Γ_d to which $q(Q) = q$ belongs. Theorem 2 shows that this construction, with q ranging over a set of representatives of the cosets in question, yields a representative of each class in the genus of f, provided $k > 4$ and f is indefinite.

9. Lower bound for the class-number (preliminary). We shall deduce Theorem 1 from Theorem 2 and

THEOREM 3. *Let S by an automorph of f with denominator prime to $d(f)$. Then $q(S)$, defined in §7, is in $\Gamma(f)$ in any case, and in $\Gamma^+(f)$ if and only if either $\Gamma(f) = \Gamma^+(f)$ or $|S| = +1$.*

It is difficult to prove this theorem directly. We shall deduce it for $|S| = +1$ from Lemma 7; and we note here that the result for $|S| = -1$ then follows on considering $U(\mathbf{z})S$, for suitable $\mathbf{z}$, and using 7.6.

Deduction of Theorem 1 *from Theorems* 2,3. We consider first assertion (ii). Suppose f has an integral automorph S with $|S| = -1$. Then obviously $q(S) = 1 \in \Gamma^+(f)$. So by Theorem 3 we have $\Gamma(f) = \Gamma^+(f)$. This shows that assertion (i) need only be proved for unrestricted equivalence.

Now consider the forms f^Q constructed in §8, with q ranging over a set of representatives of the cosets in Γ_d of $\Gamma(f)$. Theorem 1(i) follows if we prove that these forms are all inequivalent. As in the proof of Theorem 2 this can be reduced to proving that f^Q, with $q(Q) = q$, is not equivalent to f unless q is in $\Gamma(f)$. Now if $f^{QT} = f$, T integral, then QT is an automorph of f and so $q(QT) \in \Gamma(f)$, by Theorem 3. But clearly $q(QT) = q(Q)$.

Theorem 1(iii) now follows for $k > 4$ (and f indefinite), as far as unrestricted equivalence is concerned, since by Theorem 2 every form in the genus is equivalent to one of the f^Q. For proper equivalence, we modify the construction of the set of forms f^Q by making q range over a set of representatives of the cosets of $\Gamma^+(f)$. (For $k = 3$, see **3**, Theorem 1.)

10. The groups $\Gamma(p, f)$. We shall see, in Theorem 4, that Theorem 2 is in most cases (when f is indefinite) sufficient to prove that the class number is 1. Theorem 2 also tells us whether two given forms in the same genus are equivalent, provided that we can find an R by which they are related (which is not very difficult) and determine the groups $\Gamma(f)$, $\Gamma^+(f)$. In §5 we have seen how to find a q in $\Gamma(f)$ which, adjoined to $\Gamma^+(f)$, generates $\Gamma(f)$. Lemma 1 then determines $\Gamma^+(f)$, if we can determine the groups $\Gamma(p, f)$. In so doing, we shall throughout this section assume 3.3 or 3.4, with a suitable sufficiently large β.

We may thus replace the A in 4.52 by

$$[p^{\lambda_1}, \ldots, p^{\lambda_k}]$$

if $p \neq 2$, and by

$$[2^{\mu_1}, 2^{\mu_1}, \ldots, 2^{\mu_\nu}, 2^{\mu_\nu}, 2^{\lambda_{2\nu+1}}, \ldots, 2^{\lambda_k}]$$

if $p = 2$. For the matrix of a form ϕ_ρ has odd determinant, and so may be replaced by the 2×2 identity matrix without affecting 4.52; and the a_{ij}, prime to p, may be cancelled out in any case. 4.52 thus reduces for $p \neq 2$ to

$$10.1 \qquad p^\delta | p^{\lambda_i} t_i, \text{ implying } p^{\delta+1} | p^{\lambda_i} t_i^2 \text{ if } \lambda_i \neq \delta,$$

for $i = 1, \ldots, k$. And for $p = 2$ 4.52 reduces to

$$10.2 \qquad 2^\delta | 2^{\mu_\rho} t_{2\rho-1}, 2^{\mu_\rho} t_{2\rho}, \text{ implying } 2^{\delta+1} | 2^{\mu_\rho} \phi_\rho \text{ if } \mu_\rho \neq \delta,$$

for $\rho = 1, \ldots, \nu$, and

$$10.3 \qquad 2^\delta | 2^{\lambda_i+1} t_i, \text{ implying } 2^{\delta-1+|\delta-\lambda_i-1|} | 2^{\lambda_i} t_i^2,$$

for $i = 2\nu + 1, \ldots, k$.

We now prove (see **3**, Lemma 3, 598 for the case $k = 3$):

LEMMA 5. (a) *If $p \neq 2$ then $\Gamma(p, f)$ is generated by adjoining the integers*

$$a_{ij} p^{\lambda_i + \lambda_j}, \qquad i, j = 1, \ldots, k.$$

each with its square factor removed, to the group of quadratic residues given by $(v|p) = 1$ or to the group Γ_p given by $(v, p) = 1$ according as $\lambda_i = \lambda_j$ does or does not imply $i = j$.

(b) *In case $\nu \neq 0$, $\Gamma(2, f) = \Gamma$ if the exponents λ_i, μ_ρ are not all of the same parity, or if for any i, j we have*

$$10.4 \qquad \lambda_i = \lambda_j, a_i \equiv a_j \pmod 4, 2\nu < i < j \leqslant k;$$

Otherwise $\Gamma(2, f) = \Gamma_2$.

(c) *If $\nu = 0$, then $\Gamma(2, f)$ is generated by the subgroup of Γ with $v \equiv 1 \pmod 8$, together with the integers*

$$a_{ij} 2^{\lambda_i + \lambda_j},$$

each with its square factor removed, and also, if the stated conditions hold, the following integers, each with square factor removed:

(i) $1 + a_{ij}$, *if* 10.4 *holds for any i, j;*

(ii) -3, *if any two exponents differ by* 0, 2 *or* 4;

(iii) $1 + 2a_{ij}$, *if $\lambda_j - \lambda_i = 1$ or* 3;

(iv) -1, *if there are three exponents no two of which differ by more than* 3.

Proof. It is convenient to write

$$b = p^\delta b', p \nmid b';$$

and to note that we are concerned only with the parity of δ and the value of $(b'|p)$, or the residue ± 1 or ± 3 of b' modulo 8 if $p = 2$. For if $(v|p) = 1$, or $v \equiv 1 \pmod 8$, it is obviously possible to take $b_2 = b_1 v$ in 4.6; and so all such v are in $\Gamma(p, f)$.

(a) Using 10.1, we write 4.54 as

$$b' \equiv \sum_{\lambda_i = \delta} a_{ii} t_i^2 \pmod p.$$

The sum must not be empty, or $p|b'$, so δ is equal to some λ_i. If the sum contains two terms or more we can have $(b'|p) = \pm 1$ as we choose; but otherwise $(b'|p) = (a_i|p)$. Putting in 4.6 the values of b so found, we clearly obtain the stated result.

(b) We can choose t_1, t_2 so that $\phi_1(t_1, t_2)$ has any desired odd residue modulo 8; then with $t_3 = \ldots = t_k = 0$ we have $\delta = \mu_1$ and any desired b'. Similarly, δ can be taken equal to any of the μ_1, or, as shown below, to any of the λ_i. It is therefore sufficient to consider whether, if the μ_ρ, λ_i are all of the same parity, δ can be of the opposite parity. If so, then 10.2, 10.3, 4.54 give

$$10.5 \qquad \sum_{\delta-2<\lambda_i<\delta} 2^{\lambda_i} a_i t_i^2 \equiv 2^\delta \pmod{2^{\delta+1}},$$

and this sum contains only terms with $\lambda_i = \delta - 1$. This is easily seen to be impossible unless 10.4 holds.

(c) Putting $t_i = 1$ and $t_j = 0$ for $j \neq 1$, we see that we can satisfy 4.54, which by 10.2, 10.3 reduces to

$$10.6 \qquad \sum_{\delta-4<\lambda_i<\delta+2} 2^{\lambda_i} a_i t_i^2 \equiv 2^\delta b' \pmod{2^{\delta+3}},$$

with $\delta = \lambda_i$, $b' = a_i$. Thus we can have $b_1 b_2 = a_i a_j 2^{\lambda_i+\lambda_j}$ in 4.6.

If 10.4 holds, we take $i, j = 1, 2$ for convenience, and put $t_1 = 1$, $t_2 = 1$, 2, $t_3 = \ldots = t_k = 0$. 10.6 is satisfied with

$$\delta = \lambda_1 + 1, \lambda_1;\ b' = \tfrac{1}{2}(a_1 + a_2), a_1 + 4a_2.$$

Putting these values of b, and also $2^{\lambda_1} a_1$, in 4.6, we find that $\Gamma(2, f)$ contains v congruent to $1 + a_1 a_2, -3 \pmod{16}$, as asserted in clauses (i), (ii) of part (c) of the lemma; and for clause (ii) 10.42 is not needed.

To prove that $\Gamma(2, f)$ contains -3, if $\lambda_2 - \lambda_1 = 2$ or 4, or -1 or $3 \equiv 1 + 2a_1 a_2 \pmod 8$, if $\lambda_2 - \lambda_1 = 1$ or 3, write

$$\lambda_2 - \lambda_1 = 1 + \epsilon + 2\eta, \quad \epsilon = 0 \text{ or } 1, \quad \eta = 0 \text{ or } 1.$$

Put $t_1 = 2^\eta$, $t_2 = 1$, and all other $t_i = 0$. We find that 10.6 holds with $\delta = \lambda_1 + 2\eta$, $b' = a_1 + 2^{1+\epsilon} a_2$. Using this b and $2^{\lambda_1} a_1$ in 4.6, we have $v \equiv 1 + 2^{1+\epsilon} a_1 a_2 \pmod 8$ in $\Gamma(2, f)$.

Clause (iv) of part (c) is easily seen to be redundant unless the three exponents in question, which for convenience we take to be $\lambda_1, \lambda_2, \lambda_3$, are all of the same parity. If so, clause (ii) applies and we need only find a v with $v \equiv -1 \pmod 4$. This is trivial if the a_i have not all the same residue (mod 4). Taking therefore

$$\lambda_1 = \lambda_3 - 2\epsilon, \lambda_2 = \lambda_3 - 2\eta, \epsilon = 0 \text{ or } 1, \eta = 0 \text{ or } 1,$$

and

$$t_1 = 2^\epsilon, t_2 = 2^\eta, t_3 = 1, t_4 = \ldots = t_k = 0,$$

we see that 10.6 holds with

$$\delta = \lambda_3, b' = a_1 + a_2 + a_3 \equiv -a_3 \pmod 4,$$

which gives the desired result.

It remains to be seen whether we have missed any of the possibilities for $\delta \pmod 2$ or for $b' \pmod 8$. As far as δ is concerned, the argument of (b) still holds. Considering the residue of b' modulo 8, suppose first that no two of the exponents λ_i differ by 0, 2, or 4. Then if the sum in 10.6 contains three or more terms, that in 10.5 is either empty or contains a single term with exponent $\delta - 1$; in either case 10.5 is impossible. If on the other hand the sum in 10.6 contains at most two terms then the number of possibilities to be considered is very small, and it can easily be seen that b' has always a residue modulo 8 previously obtained with the same $\delta \pmod 2$.

Suppose now that there are two exponents whose difference is 0, 2, or 4. Then we have already proved $-3 \in \Gamma(2, f)$, and so need only consider $b' \pmod 4$, that is, we may reduce 10.6 $\pmod{2^{\delta+2}}$. This means, using 10.3, that the terms with exponents $\delta - 4$, $\delta + 2$, go out. We may now suppose that no two exponents differ by 1 or 3; for if such a difference occurs we already know that a v in $\Gamma(2, f)$ can be $\equiv -1 \pmod 4$, so that the residue of b' modulo 4 need not be considered. For a similar reason, we assume that no three exponents have differences all $\leqslant 2$, by (c) (iv) of the lemma. Now the number of possibilities to be considered is again very small, and we omit the remaining details.

COROLLARY. *If $p \neq 2$ and Γ_p is not included in $\Gamma(p, f)$ then $p^{\frac{1}{2}k(k-1)} | d$.*
If Γ_2 is not included in $\Gamma(2, f)$ then $4^{\frac{1}{2}k(k-1)} | d$; and if -3 is not in $\Gamma(2, f)$, then $8^{\frac{1}{2}k(k-1)} | d$.

Proof. For odd p, the present hypothesis, with part (a) of the lemma, shows that the exponents λ_i are all unequal. Their sum, say θ, is thus at least $\frac{1}{2}k(k-1)$; and obviously $p^\theta | d$.

For $p = 2$, either hypothesis, with part (b) of the lemma, gives $\nu = 0$. Now with θ as above we have

$$2^{\theta'} | d, \quad \theta' = \theta + 2\,[\tfrac{1}{2}k].$$

If $\Gamma(2, f)$ contains no $v \equiv -1 \pmod 4$, or no $v \equiv -3 \pmod 8$, then clauses (iii) and (iv), or (ii), of part (c) of the lemma show that

$$\theta \geqslant 0 + 0 + 4 + \ldots, \text{ or } \theta \geqslant 0 + 1 + 6 + \ldots .$$

By a simple calculation, this gives $\frac{1}{2}\theta'$ or $\frac{1}{2}\theta' \geqslant \frac{1}{2}k(k-1)$, which completes the proof.

We deduce:

THEOREM 4. *Suppose that f is indefinite, $k \geqslant 3$, and let d_1 be the greatest integer whose $\frac{1}{2}k(k-1)$th power divides d. Suppose also that $d_1 = 1,2,4$, p or $2p$, $p \equiv -1 \pmod 4$. Then the class-number of f, in the strict sense, is 1.*

Proof. It is sufficient, by Theorem 1(iii), to show that $\Gamma^+(f) = \Gamma_d$. We know (since n, q in 4.4 may be replaced by $-n$, $-q$) that -1 is in $\Gamma^+(f)$. So it suffices to find a subgroup of $\Gamma^+(f)$ which either coincides with Γ_d or is a

subgroup of index 2 of Γ_d, not containing -1. We obtain such a subgroup by putting $w = 1$ in 5.3, dropping the accent since f is indefinite. This subgroup is

$$\bigcap_{p|d} \{\Gamma(p,f) \cap \Gamma_p\} = \Big\{\bigcap_{2p|d_1} \Gamma(p,f)\Big\} \cap \Gamma_d,$$

by Lemma 5, Corollary. By the present hypotheses this reduces to Γ_d if $d_1 = 1$ or 2, and if $d_1 = 4$ to $\Gamma(2,f) \cap \Gamma_d$, including, by the corollary, all $q \equiv 1 \pmod 4$ in Γ_d. If $d_1 = p$ or $2p$, $p \equiv -1 \pmod 4$, then the subgroup in question is $\Gamma(p,f) \cap \Gamma_d$, with $(-1|p) = -1$, whence it does not include -1, if it is proper.

11. Factorization of automorphs. We prove:

LEMMA 6. *Every automorph S of f is expressible as a product of automorphs of the special type* 4.1; *that is, we may write*

$$S = U(\mathbf{t}_1) \dots U(\mathbf{t}_h),\ |S| = (-1)^h,\ h \geqslant 0, \tag{11.1}$$

for suitable $\mathbf{t}_i$, $i = 1, \dots, h$, with $f(\mathbf{t}_i)$ not zero. If p is odd and does not divide the denominator of S, then we may choose the $\mathbf{t}_i$ so that p does not divide the denominator of any of the $U(\mathbf{t}_i)$.

Proof. It suffices to prove the second part; the first is well known. We proceed by induction on k; for $k = 1$ the lemma is trivial since S can only be $\pm I$, and $U(\mathbf{t})$ can only be $-I$.

Consider first, for $k \geqslant 2$, the special case

$$f = a_{11}x_1 + g, \tag{11.21}$$
$$A = [2a_{11}, B], \tag{11.22}$$
$$S = [1, T], \tag{11.23}$$

T being necessarily an automorph of the $(k-1)$-ary form $g = g(x_2, \dots, x_k)$. For such f, consider the $U(\mathbf{t})$ with $t_1 = 0$; we see from 4.1 that

$$U(\mathbf{t}, A) = [1, U(\boldsymbol{\zeta}, B)], \text{ if } \mathbf{t} = \{0, \boldsymbol{\zeta}\} \tag{11.3}$$

and 11.2 holds. The inductive argument thus gives the result at once. Note that if 11.21 holds and the first column of S is $\{1,0, \dots, 0\}$, then the first row of S must be $(1,0, \dots, 0)$, so that 11.23 holds for suitable T.

Now make the weaker hypothesis that S has first column $\{1,0, \dots, 0\}$ and $p \nmid a_{11}$. The substitution

$$x_1 \to x_1 - a_{12}x_2 - \dots - a_{1k}x_k,\ x_i \to 2a_{11}x_i \qquad (i > 1),$$

has a matrix P with determinant $(2a_{11})^{k-1}$ prime to p. It takes f into a form f^P of the type 11.21, with an automorph $P^{-1}SP$ which has denominator prime to p and first column $\{1,0, \dots, 0\}$. So by 4.3 with $R = P$ we have the desired result in this less special case.

Now in the general case (using 4.3 again, with suitable integral R with determinant 1) we may suppose $p \nmid a_{11}$ (since f may be taken to be primitive).

The result will follow from what we have already proved if we can find $\mathbf{t}$ such that $U(\mathbf{t})$ has denominator prime to p and $U(\mathbf{t})S$ has first column $\{1,0,\ldots,0\}$; that is, if

$$U(\mathbf{t})S\mathbf{y} = \mathbf{y} \quad \text{where} \quad \mathbf{y} = \{1,0,\ldots,0\}.$$

For if so, we have $U(\mathbf{t})S = S_1$, $S = U^{-1}(\mathbf{t})S_1 = U(\mathbf{t})S_1$, for an S_1 for which the result has been proved. It will also suffice if we can make $U(\mathbf{t})S\mathbf{y} = -\mathbf{y}$; for the denominator of $U(\mathbf{y})$ is a divisor of $a_{11} = f(\mathbf{y})$, and so we may introduce a factor $U(\mathbf{y})$ (see 4.21).

It suffices therefore to find $\mathbf{t}$ such that

$$U(\mathbf{t})S\mathbf{y} = \pm\mathbf{y}, \mathbf{y} = \{1,0,\ldots,0\}, p \nmid f(\mathbf{t});$$

for the last of these conditions ensures that p does not divide the denominator of $U(\mathbf{t})$. We take $\mathbf{t} = \mathbf{y} \pm S\mathbf{y}$, and it suffices to prove that, with proper choice of the ambiguous sign, we have

11.41 $$U(\mathbf{y} \pm S\mathbf{y})S\mathbf{y} = \pm S\mathbf{y},$$

11.42 $$f(\mathbf{y} \pm S\mathbf{y}) \not\equiv 0 \pmod{p},$$

for $\mathbf{y}$ such that $p \nmid f(\mathbf{y})$. 11.42 is clear from

$$f(\mathbf{y} \pm S\mathbf{y}) = \tfrac{1}{2}(\mathbf{y}' \pm \mathbf{y}'S')A(\mathbf{y} \pm S\mathbf{y}) = f(\mathbf{y}) + f(S\mathbf{y}) \pm \mathbf{y}'AS\mathbf{y} = 2f(\mathbf{y}) \pm \mathbf{y}'AS\mathbf{y}.$$

For if 11.42 fails for both choices of the sign, then $p|4f(\mathbf{y})$. Now (with either sign) $U(\mathbf{y} \pm S\mathbf{y})$ takes $\mathbf{y} \pm S\mathbf{y}$ into $-(\mathbf{y} \pm S\mathbf{y})$ by 4.21, and leaves $\mathbf{y} \mp S\mathbf{y}$ invariant, by 4.22; for

$$(\mathbf{y}' \pm \mathbf{y}'S')A(\mathbf{y} \pm S\mathbf{y}) = 2f(\mathbf{y}) - 2f(S\mathbf{y}) = 0.$$

Hence 11.41 holds; and this completes the proof. It is of interest to note that we cannot always take the $U(\mathbf{t}_i)$ in 11.1 to have odd denominators when the denominator of S is odd. That is, the second part of the lemma fails for $p = 2$. To show this, take $k = 4$,

$$f = x_1^2 + x_1x_2 + x_2^2 + x_3^2 + x_3x_4 + x_4^2,$$

and let S be the matrix interchanging x_1 and x_3, x_2 and x_4. If $U(\mathbf{t})$ has odd denominator, then, by 4.1, 4.5 must hold, with $\delta = 0$ since $d = 9$ is odd. That is, $f(\mathbf{t})$ must be odd. This gives that t_1, t_2 are both even and t_3, t_4 not both even, or vice versa. Then a simple calculation shows that the (i, j) element of $U(\mathbf{t})$ must be even for $i \leqslant 2, j > 2$ or vice versa. Any product of matrices with this property has the same property; but S has not.

12. Definition and properties of $v(A, S)$. We define $v(S) = v(A, S)$ by 11.1 and

12.1 $$u^2v(S) = \prod_{i=1}^{h} f(\mathbf{t}_i),\ u \text{ integral}, v(S) \in \Gamma.$$

Although the factorization 11.1 is not unique, it is known (see **2**, Sätze 4.4,

4.5; and **4**) that $v(A, S)$ depends only on A and S. It is essentially the *spinor norm* of S as defined by Eichler **(2)**. Clearly

12.2 $$v(S_1S_2) = v(S_2S_1) = v(S_1)\cdot v(S_2).$$

From 4.3 (with $|R| \neq 0$),

12.3 $$v(R'AR, R^{-1}SR) = v(A, S).$$

In case 11.2 holds, we take the factors in 11.1 to be of the type 11.3, and so

12.4 $$v(A, S) = v(B, T).$$

The property of $v(S)$ that we need to prove Theorem 3 is given in the first assertion of the following lemma; the second assertion is put in to simplify the proof.

LEMMA 7. *Let S be any automorph of f with denominator prime to p. Let v_1 be any element of Γ such that, for suitable u, 4.5 can be satisfied with $b = u^2v_1$. Then $v(S)$ is in $\Gamma(p, f)$ if $|S| = 1$, in $v_1\cdot\Gamma(p, f)$ if $|S| = -1$.*

Proof for $p \neq 2$. First suppose $S = U(\mathbf{t})$. The hypothesis that p does not divide the denominator of S shows, using 4.1, that $\mathbf{t}$ must satisfy 4.51, 4.52, implying as we have seen 4.53, for some δ; whence $b = f(\mathbf{t})$ satisfies 4.54. Taking in 4.6 $b_1 = f(\mathbf{t})$, $b_2 = u^2v_1$, we find an element of $\Gamma(p, f)$ which is clearly $v_1\cdot v(U(\mathbf{t}))$. Hence $v(U(\mathbf{t}))$ is in $v_1\cdot\Gamma(p, f)$, since $v_1\cdot v_1 = 1$. This gives the result in the special case when $h = 1$ in 11.1; it follows generally by Lemma 6.

The case $p = 2$ is much more difficult since we cannot use Lemma 6. We shall proceed by induction on k; the case $k = 1$ is trivial as noted in the proof of Lemma 6. We shall also assume (see 5.6) that f is primitive; but an imprimitive $(k-1)$ary form may have to be considered in the induction. The argument used in the case $p \neq 2$ shows that the hypotheses and conclusion of the lemma are unaltered, except for interchange of the two cases, if S is replaced by $SU(\mathbf{t})$ or by $U(\mathbf{t})S$, the denominator of $U(\mathbf{t})$ being odd. We devote the next section to a preliminary simplification of the problem.

13. Proof of Lemma 7 for $p = 2$ (preliminary). We prove first:

LEMMA 8. *Write for brevity $\mathbf{y} = \{1, 0, \ldots, 0\}$. Then Lemma 7 (with $p = 2$) is true in the following three cases (assuming the inductive hypothesis):*

(i) $f(\mathbf{y}) \equiv 1 \pmod 2$, $2 \mid \mathbf{y}'A$, $S\mathbf{y} = \pm\mathbf{y}$;
(ii) $f(\mathbf{y}) \equiv 1 \pmod 2$, $2 \nmid \mathbf{y}'A$, $S\mathbf{y} = \pm\mathbf{y}$;
(iii) $f(\mathbf{y}) \equiv 4 \pmod 8$, $2 \nmid \mathbf{y}'A$, $S\mathbf{y} = \pm\mathbf{y}$.

Proof. (i) We begin with the still more special case 11.2. Taking $t_1 = 0$ in 4.5, we see that all the values of b that are possible with g in place of f are also possible with f; hence $\Gamma(2, g) \subseteq \Gamma(2, f)$. Moreover, the v_1 of Lemma 7

may be taken to be a value of b arising from 4.5 with $t_1 = 0$. Hence this special case can be dealt with as in Lemma 6, using 12.4, instead of factorizing the matrix T, to apply the inductive hypothesis.

In the general case, we have a_{11} odd and $a_{12}, \ldots, a_{1k}$ all even. As in the proof of Lemma 6, it suffices to remove the product terms involving x_1 by a transformation with integral coefficients and odd determinant, and then apply 12.3. The required transformation is

$$x_1 \to x - \tfrac{1}{2}a_{12}x_2 - \ldots - \tfrac{1}{2}a_{1k}x_k, \qquad x_1 \to a_{11}x_i \quad (i > 1).$$

(ii) The transformations needed to make f satisfy 3.4 can be chosen so as to leave a_{11} and $\mathbf{y}$ invariant. We may therefore assume 3.4, with $\mu_1 = 0$ since $2 \nmid \mathbf{y}'A$ means that one of $a_{12}, \ldots, a_{1k}$ (necessarily a_{12} in 3.4) is odd; and we write 3.4 for brevity as

$$13.1 \qquad f \equiv a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2 + \psi \pmod{2^\beta} \ (a_{11} \text{ odd}).$$

Write $M = [1,4, \ldots, 4]$.
It is clear that $M^{-1}SM$, which is

$$\begin{pmatrix} \pm 1 & 4\mathbf{a}' \\ 0 & S_{22} \end{pmatrix}$$

if S is

$$\begin{pmatrix} \pm 1 & \mathbf{a}' \\ 0 & S_{22} \end{pmatrix},$$

has the same odd denominator as S and satisfies $M^{-1}SM\mathbf{y} = \mathbf{y}$ (that is, has $\mathbf{y} = \{1,0, \ldots, 0\}$ as its first column). $M^{-1}SM$ is an automorph of

$$13.2 \qquad f^M \equiv a_{11}x_1^2 + 4a_{12}x_1x_2 + 16a_{22}x_2^2 + 16\psi \pmod{2^\beta}.$$

f^M satisfies the conditions of part (i) of the lemma, and so Lemma 7 is true with f^M, $M^{-1}SM$ for f, S. It follows also for f, S, using 12.3 with $R = M$, if we can prove that $\Gamma(2, f^M) = \Gamma(2, f)$.

Now from 13.2 we see that f^M goes, by a trivial unimodular transformation, into a form congruent modulo 2^β to

$$13.3 \qquad a_{11}x_1^2 + 4a_{22}'x_2^2 + 16\psi, a_{22}' \equiv -a_{11} \pmod 4.$$

$\Gamma(2, f)$ includes Γ_2 by Lemma 5(b); as does $\Gamma(2, f^M)$, by Lemma 5(b) if applicable, or by Lemma 5(c), which shows that $\Gamma(2, f^M)$ contains -3 (clause (ii)) and also an integer congruent to $a_{11}a_{22}' \equiv -1 \pmod 4$. To prove $\Gamma(2, f^M) = \Gamma(2, f)$ we therefore need only show that both or neither of these groups contains even v. Comparing 13.1 and 13.3, with ψ in each case written out in full, we see that both or neither contain two exponents of opposite parity, and both or neither contain terms satisfying 10.4. This gives the result.

(iii) This case can be reduced to case (ii). We use the same M, and cancel a divisor 4 from f^M. The argument is similar but a little simpler.

We deduce

LEMMA 9. *Lemma 7 is true for $p = 2$ (assuming the inductive hypothesis) if there exist either* $\mathbf{y}$, δ *satisfying*

$$f(\mathbf{y}) \equiv 1 \pmod 2,\ 2|\mathbf{y}'A,\ 2^\delta \| f(\mathbf{y} \pm S\mathbf{y}),\ 2^\delta \| (\mathbf{y}' \pm \mathbf{y}'S')A, \tag{13.4}$$

(with either sign) or $\mathbf{z}$ *satisfying*

$$\mathbf{z}'AS\mathbf{z} \equiv 1 \pmod 2. \tag{13.5}$$

Note that 13.4 can be satisfied, if at all, with primitive $\mathbf{y}$, so we may suppose $\mathbf{y} = \{1,0,\ldots,0\}$.

Proof. From 4.1, 13.43, 13.44 we see that the denominator of $U(\mathbf{y} \pm S\mathbf{y})$ is odd. Hence we may, as noted at the end of § 12, replace S by $U(\mathbf{y} \pm S\mathbf{y})S = S_1$, say. As we saw in the proof of Lemma 6, $S_1\mathbf{y} = \pm \mathbf{y}$. Thus by assertion (i) of Lemma 8, 13.4 implies Lemma 7.

Now assume 13.5, with $f(\mathbf{z})$ odd. We have

$$f(\mathbf{z} + S\mathbf{z}) = 2f(\mathbf{z}) + \mathbf{z}'AS\mathbf{z} \equiv 1 \pmod 2,$$

so the denominator of $U(\mathbf{z} + S\mathbf{z})$ is odd. Replacing S by $U(\mathbf{z} + S\mathbf{z})S$, Lemma 8(ii) with $\mathbf{y} = \mathbf{z}$ gives the result. If $f(\mathbf{z}) \equiv 4 \pmod 8$, we similarly apply Lemma 8 (iii).

If 13.5 holds with $f(\mathbf{z}) \equiv 2 \pmod 4$ or $0 \pmod 8$, we need only show that there exists $\mathbf{y}$ with $\mathbf{y}'AS\mathbf{y}$ odd and $f(\mathbf{y})$ odd or congruent to 4 modulo 8. We can find $\boldsymbol{\zeta}$ with $\mathbf{z}'A\boldsymbol{\zeta}$ odd, since 13.5 gives $2 \nmid \mathbf{z}'A$. We put $\mathbf{y} = \mathbf{z} + 2a\boldsymbol{\zeta}$, with suitable a. Clearly

$$\mathbf{y}'AS\mathbf{y} \equiv \mathbf{z}'AS\mathbf{z} \equiv 1 \pmod 2,$$
$$f(\mathbf{y}) = f(\mathbf{z}) + 2a\mathbf{z}'A\boldsymbol{\zeta} + 4a^2 f(\boldsymbol{\zeta}),$$

whence $f(\mathbf{y}) = f(\mathbf{z}) \pm 2, f(\mathbf{z}) + 4 \pmod 8$ for $a = \pm 1, 2$. Hence the result.

14. Completion of proof of Lemma 7; proof of Theorem 3. Suppose first that $\mu_1 = 0$ in 3.4. Then Lemma 7 is true for $p = 2$ if we can satisfy 13.5. This is possible unless AS is congruent (mod 2) to a skew matrix. We suppose therefore that this is so; and further, since we may replace S by $SU(\mathbf{t})$ if $f(\mathbf{t})$ is odd, making the denominator of $U(\mathbf{t})$ odd, that $ASU(\mathbf{t})$ has the same property for every such $\mathbf{t}$. We shall show that this leads to a contradiction by assuming 3.4, with $\mu_1 = 0$ and a_{11} odd, as we may since ϕ_1 represents odd integers. We take $\mathbf{t} = \{1,0,\ldots,0\}$, so that $f(\mathbf{t}) = a_{11}$ is odd. A simple calculation shows that $U(\mathbf{t})$ is congruent (mod 2) to the matrix with 1's on its diagonal and in the (1, 2) position and 0's elsewhere. Then if $\mathbf{s}_1$, $\mathbf{s}_2$ are the first two column vectors of S, those of AS are, by 3.4, congruent modulo 2 to $\mathbf{s}_2$, $\mathbf{s}_1$; and those of $ASU(\mathbf{t})$ to $\mathbf{s}_2, \mathbf{s}_1 + \mathbf{s}_2$. With AS and $ASU(\mathbf{t})$ both skew modulo 2 we must have $\mathbf{s}_2 \equiv 0$, $|S| \equiv 0 \pmod 2$, which is impossible.

We may therefore suppose, since f is assumed primitive, that in 3.4 we have $\lambda_1 = 0$ and all the μ_ρ positive. We may also suppose that no μ_ρ is 1, since otherwise Lemma 5(b) gives $\Gamma(2, f) = \Gamma$, and we have nothing to prove.

If three or more of the λ_t are 0, then 10.4 holds, and by Lemma 5(b) or (c)(i), (ii), (iv), we again have $\Gamma(2, f) = \Gamma$. We therefore assume that at most two of the λ_t vanish, and so rearranging the terms of 3.4 we may suppose

$$A \equiv [2, 2^{\lambda_2+1}, 0, \ldots, 0] \pmod 4. \tag{14.1}$$

It suffices now to deduce from 14.1 that 13.4 can be satisfied with $\mathbf{y} = \{1, 0, \ldots, 0\}$. This choice of $\mathbf{y}$ certainly satisfies 13.41 and 13.42. We choose the sign in 13.43 so that this condition holds with $\delta = 1$ or 2; for the sum of the two numbers

$$f(\mathbf{y} \pm S\mathbf{y}) = 2f(\mathbf{y}) + \mathbf{y}'AS\mathbf{y} \text{ is } 4f(\mathbf{y}) \equiv 4 \pmod 8.$$

Now 13.44 is certainly satisfied if $\delta = 1$. 13.43 holds with $\delta = 1$ if $\mathbf{y}'AS\mathbf{y} \equiv 0 \pmod 4$. If this is not so, then with $S\mathbf{y} = \boldsymbol{\xi} = \{\xi_1, \ldots, \xi_k\}$ we have ξ_1 odd. If so, then by 14.1 we have

$$1 \equiv f(\mathbf{y}) \equiv f(S\mathbf{y}) \equiv 1 + 2^{\lambda_2}\xi_2 \pmod 2.$$

Thus $2^{\lambda_2}\xi_2$ is even, and this with 14.1 gives 13.44 with $\delta = 2$.

Proof of Theorem 3. For the reason noted in §9, we may suppose $|S| = 1$. Since no p dividing d divides the denominator of S, we have $v(S) \in \Gamma(p, f)$ for each such p, by Lemma 7. If f is definite, then by 11.1 with h even since $|S| = 1$, we have $v(S) > 0$. Hence if we write

$$v(S) = wq_1, \; w|d, \; q_1 > 0, \; q_1 \in \Gamma_d,$$

we have by Lemma 1 $q_1 \in \Gamma^+(f)$. It suffices to prove $q_1 = q(S)$.

This is equivalent to showing that if $p \nmid d$ then $p|v(S)$ if and only if $p|q(S)$. We shall prove this for all S with denominators prime to d (without the restriction $|S| = 1$). For simplicity we shall assume that either p does not divide the denominator of S, or the numbers $\theta_t(S,p)$ defined in §7 are $-1, 0, \ldots, 0, 1$ (see 7.3, with $R = S$). It is clear from §9 that these are the only cases of Theorem 3 that are needed to prove Theorem 1. Other cases can however be dealt with similarly. In the second case, Lemma 4 shows that we can write $S = U(\mathbf{z})S_1$, where $p \| f(\mathbf{z})$ and S_1 has denominator prime to p.

Now in the first case, when p does not divide the denominator of S, we have from Lemma 7, $v(S) \in \Gamma(p, f)$; that is, by 4.7, $p \nmid v(S)$, and clearly $p \nmid q(S)$. In the other case p divides $q(S)$, and using 12.2 we have $v(S) = v(U(\mathbf{z}))$. $v(S_1)$, that is, $v(S)$ is $f(\mathbf{z})v(S_1)$ with its square factor removed. Since $p \| f(z)$ and, by what we have just proved, $p \nmid v(S_1)$, this gives $p|v(S)$, and the proof is complete.

References

1. H. Brandt, *Über quadratische Kern- und Stamm-formen*, Festschrift zum 60. Geburstag von Professor Andreas Speiser (Zürich, 1945).
2. M. Eichler, *Quadratische Formen und orthogonale Gruppen* (Berlin, 1952).
3. B. W. Jones and G. L. Watson, *On indefinite ternary quadratic forms*, Can. J. Math., *8* (1956), 592–608.
4. R. Lipschitz, *Untersuchungen über die Summen von Quadraten* (Bonn, 1886).
5. A. Meyer, *Ueber indefinite quadratische Formen*, Viertelschr. Naturforsch. Ges. Zürich, *36* (1891), 241–250.
6. G. Pall, *On the order invariants of integral quadratic forms*, Quart. J. Math. (Oxford), *6* (1935), 30–51.
7. G. L. Watson, *Representation of integers by indefinite quadratic forms*, Mathematika, *2* (1955), 32–38.

University College,
London

CONGRUENCES FOR THE COEFFICIENTS OF MODULAR FORMS AND SOME NEW CONGRUENCES FOR THE PARTITION FUNCTION

MORRIS NEWMAN

If n is a non-negative integer, define $p_r(n)$ as the coefficient of x^n in

$$\prod_{n=1}^{\infty} (1 - x^n)^r;$$

otherwise define $p_r(n)$ as 0. In a recent paper **(2)** the author established the following congruence:

Let $r = 4, 6, 8, 10, 14, 26$. Let p be a prime greater than 3 such that $r(p + 1)/24$ is an integer, and set $\Delta = r(p^2 - 1)/24$. Then if $R \equiv r \pmod{p}$ and $n \equiv \Delta \pmod{p}$,

$$p_R(n) \equiv 0 \pmod{p}. \tag{1}$$

The choices $r = 4$, $p = 5$; $r = 6$, $p = 7$; and $r = 10$, $p = 11$ (all with $R = -1$) give the Ramanujan congruences

$$p(5n+4) \equiv 0 \pmod{5} \tag{2}$$

$$p(7n+5) \equiv 0 \pmod{7} \tag{3}$$

$$p(11n+6) \equiv 0 \pmod{11}. \tag{4}$$

It is also possible to determine from (1) the Ramanujan congruence modulo 25.

In this paper we establish a congruence similar to (1) (Theorem 1). By appropriate specialization we obtain congruences of the Ramanujan type for $p(n)$ with modulus 13 (formulas (11), (12) and (14)). A significant difference emerges, however. The congruences (2), (3), (4) are statements concerning arithmetic progressions. The congruence (14) is valid for sequences which are essentially geometric progressions. Thus divisibility of $p(n)$ by 13 seems a rarer phenomenon than divisibility by 5, 7 or 11.

The author has shown **(3; 4)** that the following identity is valid:

Suppose that r is even, $0 < r \leqslant 24$. Let p be a prime greater than 3 such that $\delta = r(p - 1)/24$ is an integer. Then for all integral n

$$p_r(np + \delta) = p_r(n)p_r(\delta) - p^{\frac{1}{2}r-1}p_r\left(\frac{n - \delta}{p}\right). \tag{5}$$

Thus for $r \geqslant 4$,

$$p_r(np + \delta) \equiv p_r(n)p_r(\delta) \pmod{p}. \tag{6}$$

Received April 1, 1957. The preparation of this paper was supported (in part) by the Office of Naval Research.

We use this congruence to prove the following theorem:

THEOREM 1. Suppose that r is even, $4 \leqslant r \leqslant 24$. Let p be a prime greater than 3 such that $\delta = r(p-1)/24$ is an integer. Then if Q,n are integers and $R = Qp + r$,

$$p_R(np + \delta) \equiv p_r(\delta)p_{Q+r}(n) \pmod{p}. \tag{7}$$

Proof. We have

$$\sum_{n=0}^{\infty} p_R(n)x^n = \prod_{n=1}^{\infty} (1-x^n)^R = \prod_{n=1}^{\infty} (1-x^n)^{Qp+r}$$
$$\equiv \prod_{n=1}^{\infty} (1-x^{np})^Q \prod_{n=1}^{\infty} (1-x^n)^r \pmod{p}.$$

Thus comparing coefficients,

$$p_R(n) \equiv \sum_{0<j\leqslant n/p} p_Q(j)p_r(n-pj) \pmod{p}.$$

Replace n by $np + \delta$. Since $\delta/p < 1$, j runs from 0 to n inclusive, and making use of (6) we obtain

$$p_R(np + \delta) \equiv p_r(\delta) \sum_{j=0}^{n} p_Q(j)p_r(n-j) \equiv p_r(\delta)p_{Q+r}(n) \pmod{p},$$

which proves the theorem.

We now choose $p = 13$, $r = 12$, $R = -1$, $Q = -1$. Then (7) becomes

$$p(13n + 6) \equiv p_{12}(6)p_{11}(n) \equiv 11p_{11}(n) \pmod{13}. \tag{8}$$

Formula (8) requires a knowledge of $p_{12}(6)$, which may be found in **(5)**. This congruence seems first to have been given by Zuckerman **(7)**.

Similarly the choices $p = 13$, $r = 24$, $R = 11$, $Q = -1$; and $p = 13$, $r = 10$, $R = 23$, $Q = 1$ yield

$$p_{11}(13n + 12) \equiv p_{24}(12)p_{23}(n) \equiv 8p_{23}(n) \pmod{13} \tag{9}$$
$$p_{23}(13n + 5) \equiv p_{10}(5)p_{11}(n) \equiv 4p_{11}(n) \pmod{13}. \tag{10}$$

The number $p_{10}(5)$ is also given in **(5)**, and $p_{24}(12) = \tau(13)$ may be found, for example, in Watson's table of the τ-function **(6)**.

Congruences (8), (9), (10) may now be combined in obvious fashion to give a congruence involving $p(n)$ only.

THEOREM 2. For $n \equiv 6 \pmod{13}$,

$$p(13^2n - 7) \equiv 6p(n) \pmod{13}. \tag{11}$$

It is plain that this theorem implies that $p(n)$ is divisible by 13 infinitely often, since for example $p(84)$ is divisible by 13 and $84 = 13.6 + 6$. More precisely, define the sequence

$$t_0 = 13t + 6,$$
$$t_n = 13^2t_{n-1} - 7, \quad n \geqslant 1.$$

Here t is an arbitrary integer. Replacing n by t_{n-1} in (11) we find

$$p(t_n) \equiv 6p(t_{n-1}) \pmod{13}$$

which upon iteration becomes

$$p(t_n) \equiv 6^n p(t_0) \pmod{13}.$$

We thus obtain

COROLLARY 1. Let t be an arbitrary integer. Put $a = 24t + 11$, $b = 13t + 6$. Let n be a non-negative integer and put

$$\Delta_n = \frac{13}{24}(13^{2n} - 1)$$

(Δ_n is an integer).

Then

$$p(a\Delta_n + b) \equiv 6^n p(b) \pmod{13}. \tag{12}$$

We are interested in determining when $p(b) \equiv 0 \pmod{13}$. Since $b \equiv 6 \pmod{13}$ and $p_{11}(n)$ is tabulated in **(5)**, congruence (8) may be used to determine the first few b's such that $p(b)$ is divisible by 13. We find in fact that this occurs for the following values:

(13)

t	a	b	t	a	b
6	155	84	57	1379	747
10	251	136	68	1643	890
17	419	227	69	1667	903
18	443	240	74	1787	968
24	587	318	90	2171	1176
27	659	357	95	2291	1241

We obtain therefore

COROLLARY 2. Let a, b have the values given in table (13). Then

$$p(a\Delta_n + b) \equiv 0 \pmod{13}. \tag{14}$$

Formula (14) is a new congruence of the Ramanujan type.

The size of the numbers Δ_n prevents checking (11), (12) or (14) by any existing table of $p(n)$ (Gupta's table of the partition function extends only to $n = 600$). Since the numbers t_n are all congruent to 6 modulo 13, formula (8) may be applied to the table of $p_{11}(n)$ for $1 \leqslant n \leqslant 800$ given in **(5)**. It was found that (11), (12) and (14) were verified for all values obtainable from this table.

The divisibility of $p(b)$ by 13 for the first 6 values of b given in table (13) may be checked by Gupta's table **(1)** which gives the residues of $p(n)$ modulo 13 and modulo 19 for $0 \leqslant n \leqslant 721$.

References

1. H. Gupta, *On a conjecture of Ramanujan*, Proc. Ind. Acad. Sciences A, *4*, (1936), 625–629.
2. M. Newman, *Some Theorems about* $p_r(n)$, Can. J. Math., *9* (1957). 68–70.
3. ———, *The coefficients of certain infinite products*, Proc. Amer. Math. Soc., *4*, (1953), 435–439.
4. ———, *Remarks on some modular identities*, Trans. Amer. Math. Soc., *73* (1952), 313–320.
5. ———, *A table of the coefficients of the powers of* $\eta(\tau)$, Proc. Kon. Nederl. Akad. Wetensch. Ser. A57 = Indagationes Math., *18* (1956), 204–216.
6. G. N. Watson, *A table of Ramanujan's function* $\tau(n)$, Proc. London Math. Soc., *51* (1949), 1–13.
7. H. Zuckerman, *Identities analagous to Ramanujan's identities involving the partition function*, Duke Math. J., *5* (1939), 88–110.

National Bureau of Standards
Washington, D.C.

ON CAYLEY'S PARAMETERIZATION

M. H. PEARL

1. Introduction. A matrix P with elements from an arbitrary field $\mathfrak{F}$ is called a *cogredient automorph* (c.a.) of a symmetric matrix A if $P'AP = A$, where P' is the transpose of P. A fundamental theorem concerning cogredient automorphs is:

THEOREM (Cayley). *If A is a non-singular symmetric matrix and if Q is a skew-symmetric matrix such that $A + Q$ is non-singular, then*

$$P = (A + Q)^{-1}(A - Q) \tag{1}$$

is a c.a. *of A and $I + P$ is non-singular.*

Conversely, if P is a c.a. *of A such that $I + P$ is non-singular, then there exists a unique skew-symmetric matrix Q such that P can be expressed by means of equation* (1).

The main purpose of this paper is to demonstrate the following generalization of Cayley's theorem as applied to the real field. (Henceforth all matrices are assumed to be real unless otherwise stated.)

THEOREM 1. *If A is a (not necessarily non-singular) symmetric matrix and if Q is a skew-symmetric matrix such that $A + Q$ is non-singular, then equation* (1) *defines a* c.a. *P of A whose determinant is $+1$ and having the property that A and $I + P$ span the same row space.*

Conversely, if P is a c.a. *of A whose determinant is $+1$ and if P has the property that $I + P$ and A span the same row space, then there exists a skew-symmetric matrix Q such that P is given by equation* (1).

The matrix Q is not unique. However, the size of the family of matrices Q which yield a particular c.a. P of A will be found and a set of necessary and sufficient conditions for two skew-symmetric matrices to yield the same c.a. will be given. A simple example will be included to show that Theorem 1 is false over a field of characteristic two.

2. Proof of the theorem. The first part of the theorem is immediate. Let $A + Q = U$, $A - Q = V$. Then (see **2**) $A = \frac{1}{2}(U + V)$ and

$$P'AP = \tfrac{1}{2}\, UV^{-1}(U + V)U^{-1}\, V = A.$$

Received February 6, 1957; presented to the American Mathematical Society, October 22, 1955. This paper constitutes a portion of a thesis submitted to the University of Wisconsin for the Ph.D. degree. The author wishes to thank Professor C. C. MacDuffee for his advice and encouragement.

The author also wishes to thank the referee for suggestions which shortened the proof of Theorem 4.

Furthermore $|P| = |(A+Q)^{-1}||(A-Q)| = |(A+Q)^{-1}||(A+Q)'| = +1$ and $I+P = (A+Q)^{-1}(2A)$. Thus $I+P$ and A span the same row spaces.

Proof of the converse. In order to facilitate the construction of a skew-symmetric matrix Q satisfying equation (1), we shall first simplify the forms of P and A. This can be done by repeated application of the following lemma.

LEMMA 1. *Let U be an orthogonal matrix. Then $U'PU$ is a* c.a. *of $U'AU$ if and only if P is a* c.a. *of A. Equation* (1) *holds if and only if*

$$U'PU = (U'AU + U'QU)^{-1}(U'AU - U'QU). \tag{2}$$

Moreover, $|P| = |U'PU|$ and $I+P$ spans the same row space as A if and only if $I+U'PU$ spans the same row space as $U'AU$.

When it is convenient to do so, we shall specify U and replace P, A and Q by $U'PU$, $U'AU$ and $U'QU$ respectively. In an effort to keep the notation as simple as possible, we shall refer to $U'PU$, $U'AU$ and $U'QU$ simply as P, A and Q whenever it is clear from the context what these symbols mean.

Let A be an arbitrary symmetric matrix of order n and rank r. Since there is an orthogonal matrix U such that

$$U'AU = \begin{bmatrix} \lambda_1 & & & & & & \\ & \lambda_2 & & & & & \\ & & \lambda_r & & & & \\ & & & \ddots & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix} \qquad (\lambda_i \neq 0),$$

and all the remaining elements are 0, we shall apply Lemma 1 and assume that A is in this form.

Equation (1) is equivalent to the two conditions

$$Q(I+P) = A(I-P) \tag{3'}$$

$$|A+Q| \neq 0. \tag{3''}$$

Let P and A be partitioned as follows:

$$P = \begin{bmatrix} B & E \\ C & F \end{bmatrix}, \quad A = \begin{bmatrix} d & 0 \\ 0 & 0 \end{bmatrix},$$

where B and d are of order r. Since $I+P$ spans the same row space as A, we must have $E=0$, $F=-I_1$ (where I_1 is the identity matrix of order $n-r$) and the rank of the n by r array consisting of the first r columns of $I+P$ must be r. Furthermore, since P is a c.a. of A, $B'dB = d$; also $|P| = +1$ implies that $|B| = |F|$. Equation (3') has become

$$Q\begin{bmatrix} I_2+B & 0 \\ C & 0 \end{bmatrix} = \begin{bmatrix} d(I_2-B) & 0 \\ 0 & 0 \end{bmatrix}, \tag{4}$$

where I_2 is the identity matrix of order r.

Let the rank of $I_2 + B$ be s. Since the rank of $I + P$ is r, there is a rearrangement of the rows of $I_2 + B$ and of C such that the matrix formed by the last s rows of $I_2 + B$ and the first $r - s$ rows of C is non-singular. This rearrangement can be carried out using Lemma 1 without disturbing the form of A, as there are orthogonal matrices u and v of orders r and $n - r$ respectively, whose rows are permutations of the rows of the identity matrices I_2 and I_1 and which effect the desired rearrangements when operating on $I_2 + B$ and C respectively on the left. Let $U = u \dotplus v$. After applying Lemma 1 once again, and denoting $u'Bu$ by B, $u'du$ by d and $v'Cu$ by C, equations (3′), (3″) and (4) remain unchanged. It is to be noted here for subsequent use that the set of principal submatrices of $I_2 + B$ is invariant under a similarity transformation by $u \dotplus v$.

Now partition $I + P$ into

$$\begin{bmatrix} I_2 + B & 0 \\ C & 0 \end{bmatrix} = \begin{bmatrix} G & 0 \\ H & 0 \\ G_1 & 0 \end{bmatrix},$$

where H is the non-singular matrix constructed above.

By two transformations similar to those described by Lemma 1, G and G_1 may be eliminated. It is possible to eliminate G_1 without disturbing the right side of equation (3′), for there is a matrix

$$V = \begin{bmatrix} I_3 & 0 \\ V_1 & I_4 \end{bmatrix},$$

where I_3 and I_4 are identity matrices of orders $2r - s$ and $n - 2r + s$ respectively, such that

$$V(I + P) = \begin{bmatrix} G & 0 \\ H & 0 \\ 0 & 0 \end{bmatrix}.$$

Clearly, $2r - s \geqslant r$ and hence $(V')^{-1} A = A$. Thus equation (3′) becomes

$$(V')^{-1} QV^{-1} \cdot V(I + P) = (V')^{-1} A(I - P) = A(I - P).$$

This process is repeated once again to eliminate G and, at the same time, to replace H by an I_2 which is more conveniently positioned. Let I_5 denote the identity matrix of order $r - s$ and define

$$M = \begin{bmatrix} 0 & H^{-1} & 0 \\ I_5 & -GH^{-1} & 0 \\ 0 & 0 & I_4 \end{bmatrix}.$$

Then $MV(I + P) = I_2 \dotplus 0$ and so we have

$$(5) \quad ((MV)')^{-1} Q (MV)^{-1} \cdot (MV)(I + P) = Q_1(I_2 \dotplus 0) = (M')^{-1} A(I - P)$$

where $Q_1 = ((MV)')^{-1} Q (MV)^{-1}$. A direct computation shows that

$$(M')^{-1}A(I-P)=\begin{bmatrix}(I_2+B)'d(I_2-B) & 0\\ K & 0\\ 0 & 0\end{bmatrix},$$

where the $r-s$ by r array K consists of the first $r-s$ rows of $d(I_2-B)$. Since B is a c.a. of d, $(I_2+B)'d(I_2-B)$ is skew-symmetric.

The problem has now been reduced to the construction of a skew-symmetric matrix Q which satisfies the conditions (3″) and (5). Equation (5) uniquely defines the first r rows and the first r columns of Q_1 but places no further restrictions on it. Hence, if such a matrix Q_1 exists, it must be of the form

$$\begin{bmatrix}(I_2+B)'d(I_2-B) & -K' & 0\\ K & X & -Y'\\ 0 & Y & Z\end{bmatrix}$$

and it only remains to find matrices X, Y and Z satisfying the two conditions:

(i) X and Z are skew-symmetric matrices of orders $r-s$ and $n-2r+s$ respectively,

(ii) $|A_1+Q_1|\neq 0$, where A_1 is defined to be $((MV)')^{-1}A(MV)^{-1}$. By simplifying A_1+Q_1, it will be shown that X and Y are completely arbitrary (except for the restriction that X is skew-symmetric) but that Z must also be non-singular. A computation shows that

$$A_1+Q_1=\begin{bmatrix}2(I_2+B)'d & K_1' & 0\\ 2[\delta\ \ 0] & \delta+X & -Y'\\ 0 & Y & Z\end{bmatrix},$$

where δ is the uppermost principal submatrix of order $r-s$ of d and $[\delta\ 0]$ is the $r-s$ by r array

$$\begin{bmatrix}\lambda_1 & 0 & 0 & \cdot & \cdot & \cdot & 0\\ 0 & \lambda_2 & 0 & \cdot & \cdot & \cdot & 0\\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot\\ 0 & 0 & 0 & \lambda_{r-s} & \cdot & \cdot & 0\end{bmatrix}$$

and where K_1 consists of the first $r-s$ rows of $2dB$, i.e., K_1' consists of the first $r-s$ columns of $2B'd$. By using a series of elementary transformations, it can be shown that A_1+Q_1 is equivalent to

$$\begin{bmatrix}0 & L_1 & -2\delta & 0\\ 0 & L' & 0 & 0\\ 2\delta & 0 & -\delta+X & -Y'\\ 0 & 0 & Y & Z\end{bmatrix}$$

where L is the lower right-hand principal submatrix of $2d(I_2+B)$ of order s. It is not necessary to define L_1 explicity.

By the Laplace development of the determinant, we have $|A_1+Q_1|=\pm|2\delta|^2|L'||Z|$. It is clear that $|2\delta|\neq 0$ and hence the proof of the theorem will be complete when it is shown that

(6′) the order of Z is even,

(6″) $|L| \neq 0$.

Condition (6′) follows directly from

LEMMA 2. *Let B be a* c.a. *of the non-singular matrix d and let the multiplicity of -1 as a root of B be α. Then $|B| = (-1)^{\alpha}$.*

Since B is a c.a. of d, $B = d^{-1}(B')^{-1}d$ and $xI - B = d^{-1}(xI - (B')^{-1})d$, that is, B and $(B')^{-1}$ have the same characteristic equations and hence the same characteristic roots. Thus, the characteristic roots of B, other than $+1$ and -1 occur in reciprocal pairs. Since $|B|$ is a product of these roots, the lemma follows.

Let us return to condition (6′). The order of $F = -I_1$ is $n - r$ and $|F| = (-1)^{n-r}$. Furthermore, -1 appears as a root of B with multiplicity $r - s$ and hence, by Lemma 2, $|B| = (-1)^{r-s}$. Moreover, it has been shown that $|P| = |F| \cdot |B| = +1$ and so

$$(n - r) + (r - s) = n - s$$

is even. The order of Z is

$$n - 2r + s = n - s - 2(r - s)$$

and thus is also even. We have shown that a non-singular skew-symmetric matrix Z always exists.

It remains to show that condition (6″) is always satisfied and this will constitute the second part of the proof.

3. Pr and CPr matrices. It is now possible to prove a corollary to the first half of the proof of the converse of Theorem 1 which will be used as a lemma to the second half.

The first application of Lemma 1 transformed A into $d \dotplus 0$. It is not necessary to determine what effect it had on B. However, the second application of Lemma 1, using $U = u \dotplus v$, has the property that it leaves the set of principal submatrices of $I_2 + B$ invariant. Thus, once A has been reduced to the form $d \dotplus 0$, the set of principal submatrices is fixed. We selected an arbitrary set of linearly independent rows of $I_2 + B$ and then showed that, for the given c.a. P of A, a skew-symmetric matrix Q satisfying conditions (3′) and (3″) can be found if and only if the principal submatrix of these rows, which has been denoted by L, is non-singular; that is, the non-singularity of L is independent of the particular set of rows of $I_2 + B$ selected. Furthermore, if B is a c.a. of d, there is some n for which a c.a. P of A exists which satisfies the hypotheses of the theorem and which is in the form

$$\begin{bmatrix} B & 0 \\ C & -I_1 \end{bmatrix};$$

that is, this discussion pertains to all B. Thus, we have proved part (a) of the

COROLLARY. (a) *Let B be a* c.a. *of a non-singular diagonal matrix d of order r. Let $I_2 + B$ have rank s and let $\mathbf{X}_1$ be a set of s linearly independent rows of $I_2 + B$ such that the principal submatrix of $I_2 + B$ determined by these rows is non-singular. If $\mathbf{X}_2$ is any set of s linearly independent rows of $I_2 + B$, then the principal submatrix of $I_2 + B$ determined by these rows is non-singular.*

(b) *Let b be a* c.a. *of d. If $\mathbf{Y}$ is a set of s linearly independent rows of $b^{-1}(I_2 + B)b$, then the principal submatrix determined by these rows is non-singular.*

To prove part (b), we define $B_1 = b \dotplus I$. The matrix P has a parameterization in the form of equation (1) if and only if $P_1 = B_1^{-1} PB_1$ has such a parameterization, for $P_1 = B_1^{-1}(A + Q)^{-1}(B_1')^{-1}(B_1')(A - Q)B_1 = (A + Q_1)^{-1}(A - Q_1)$, where $Q_1 = B_1'QB_1$. Moreover, $b^{-1}(I_2 + B)b = I_2 + b^{-1}Bb$ in $I + P_1$ corresponds to $I_2 + B$ in $I + P$ and thus the principal submatrix of a set of s linearly independent rows is non-singular in one if and only if it is non-singular in the other.

Schwerdtfeger **(1)** has called a matrix of rank r which has a principal non-singular submatrix of order r a Pr matrix. We shall define a CPr matrix to be a matrix of rank r with the following property: whenever a set of s rows is linearly independent, then the set of the corresponding s columns is also linearly independent and conversely; that is, the same set of rows of the transpose of the matrix is linearly independent. Equivalently, a CPr matrix can be defined as a matrix of rank r such that the principal submatrix determined by any set of r linearly independent rows is non-singular. Clearly, a CPr matrix is always a Pr matrix. The preceding corollary asserts that if B is a c.a. of a non-singular diagonal matrix d and if $I_2 + B$ is a Pr matrix, then $I_2 + B$ is a CPr matrix. Theorem 1 will follow when we have proved

THEOREM 2. *If B is a* c.a. *of a non-singular diagonal matrix d, then $I_2 + B$ is a Pr matrix.*

It is sufficient to prove this theorem for the case where the non-zero elements of d are each $+1$ or -1, since there is a non-singular diagonal matrix f such that fdf is a diagonal matrix whose diagonal elements are each $+1$ or -1. Then $f^{-1}Bf$ is a c.a. of fdf and $I_2 + f^{-1}Bf$ is a Pr matrix if and only if $I_2 + B$ is. Hence, for the remainder of the proof we can assume that d is already in this form.

Williamson **(3)** has called a c.a. of such a matrix $\mathfrak{d}$, a quasi-unitary matrix and he has given a comprehensive discussion of the problem of reducing a quasi-unitary matrix to a canonical form by a quasi-unitary similarity transformation. He has shown that, with at most an interchange of the rows and the corresponding columns, B can be made quasi-unitarily similar to a matrix of the form

$$A_0 \dotplus A_1 \dotplus \ldots \dotplus A_k \dotplus A_{k+1} \dotplus \ldots \dotplus A_{k+m},$$

where no root of A_0 is -1, where $A_1, \ldots, A_k$ are each of odd order (say the order of A_h is $2a_h + 1$, $h = 1, 2, \ldots, k$) and A_h $(1 \leqslant h \leqslant k)$ has

$$(\lambda + 1)^{2a_{k+1}}$$

as its only elementary divisor, and where $A_{k+1}, \ldots, A_{k+m}$ are each of order divisible by 4 (say the order of A_{k+h} is $4b_{k+h}$; $h = 1,2, \ldots, m$) and A_{k+h} $(1 \leqslant h \leqslant m)$ has

$$(\lambda + 1)^{2b_{k+h}}$$

as an elementary divisor of multiplicity two. By the above corollary, the property of being a *Pr* matrix is invariant under such a transformation. Let I_h be the identity matrix whose order is equal to that of A_h. This transformation does not effect the identity matrix and since $I_2 + B$ is a *Pr* matrix if and only if $I_h + A_h$ is a *Pr* matrix for each h, we may consider each $I_h + A_h$ separately.

Case I: $I_0 + A_0$. Since A_0 does not have -1 as a characteristic root, $I_0 + A_0$ is non-singular and hence is a *Pr* matrix.

Case II: $I_h + A_h$, $(1 \leqslant h \leqslant k)$. For convenience we shall drop the subscript h from I, A and a. Since $I + A$ has nullity 1, we wish to show that $I + A$ has a principal non-singular submatrix of order $a - 1$. Let W be the matrix of the same order as A which has 1's just above the main diagonal and zeros elsewhere, that is, $W = [\delta_{i,j-1}]$.[2] Then the Jordan form for A is $-I - W$. In particular, Williamson has shown that there exist matrices D and T such that

$$(7) \qquad TDT' = e \begin{bmatrix} & & & 1 \\ & & -1 & \\ & & \iddots & \\ & S & & \\ & & (-1)^e & \\ & & \iddots & \\ & -1 & & \\ & S' & & \\ 1 & & & \end{bmatrix} = e\Delta \quad (e = \pm 1)$$

and $T^{-1}(-I - W)T = A$, where D is a matrix having the same form as d (a diagonal matrix whose diagonal elements are each $+1$ or -1) and where S represents a triangular array of terms which need not be specified here since it is soon to be eliminated. Furthermore, Williamson has shown that if T is any matrix satisfying equation (7) and if $\alpha = T^{-1}(-I - W)T$, then α is quasi-unitarily similar to A. Clearly α can be considered here instead of A.

Rewrite equation (7) as $T^{-1}(e\Delta)(T')^{-1} = D$. We shall construct a matrix T satisfying this form of equation (7) and then show that the resulting α is a *Pr* matrix. First, define H to be the matrix which has $+1$'s and -1's alternating along its skew-diagonal and zeros elsewhere, that is,

$$H = [(-1)^{i+1}\delta_{i,\,2a-i+2}].$$

[2] δ is the Kronecker delta.

Then the arrays S and S' may be eliminated as follows: there exists a matrix $T_1 = \tau \dotplus I_a$, where τ is a triangular matrix of order $a+1$ which has 1's on its main diagonal and zeros above it and where I_a is the identity matrix of order a, such that $T_1^{-1}(e\Delta)(T_1')^{-1} = eH$. Now, let E be the matrix of order a which has 1's on its skew diagonal and zeros elsewhere, that is, $E = [\delta_{i,a-i+1}]$. If we now define T_2 to be

$$\begin{bmatrix} \frac{1}{2}\sqrt{2}\, I_a & 0 & -\frac{1}{2}\sqrt{2}\, E \\ 0 & 1 & 0 \\ \frac{1}{2}\sqrt{2}\, E & 0 & \frac{1}{2}\sqrt{2}\, I_a \end{bmatrix}$$

and define T to be T_1T_2, then $T^{-1}(e\Delta)(T')^{-1}$ is in the desired form, namely eD. Furthermore,

$$I + \alpha = I + T^{-1}(-I - W)T = -T^{-1}WT.$$

A computation will show that the first row of T is $[\frac{1}{2}\sqrt{2}\ 0 \ldots 0 - \frac{1}{2}\sqrt{2}]$ and the last row is $[\frac{1}{2}\sqrt{2}\ 0 \ldots 0\ \frac{1}{2}\sqrt{2}]$. Hence, if the first row and the last column of T are removed, leaving the matrix t_1 of order $2a$, then t_1 is non-singular since $|T| = \sqrt{2}|t_1| \neq 0$. Similarly, if the last row and the last column of T^{-1} are removed, leaving the matrix t_2 of order $2a$, then t_2 is non-singular since $|T^{-1}| = \sqrt{2}|t_2| \neq 0$. The principal submatrix of order $2a$ of $I + \alpha$ which is formed by removing the last row and the last column of $I + \alpha$ is $-t_2t_1$, which is non-singular. Thus, we have shown that $I + \alpha$ and hence $I + A$, are Pr matrices.

Case III: $I + A_{k+h}$, $(1 \leqslant h \leqslant m)$. As before, we shall drop the subscript $k + h$ from I, A and b. Let I_b denote the identity matrix of order $2b$. Williamson has shown that in this case there exist matrices D and T such that

$$TDT' = \begin{bmatrix} 0 & I_b \\ I_b & 0 \end{bmatrix} \tag{8}$$

and $T^{-1}((-I_b - W) \dotplus (-I_b - W')^{-1})T = A$, where D is of the same form as in Case II. Again, if T satisfies equation (8) and if

$$\alpha = T^{-1}((-I_b - W) \dotplus (-I_b - W')^{-1})T,$$

then α is quasi-unitarily similar to A. Set $V = (-I_b - W')^{-1} + I_b$. It is easily seen that T may be taken as

$$\frac{1}{\sqrt{2}}\begin{bmatrix} I_b & I_b \\ I_b & -I_b \end{bmatrix},$$

in which case

$$I + \alpha = -\tfrac{1}{2}\begin{bmatrix} W - V & W + V \\ W + V & W - V \end{bmatrix}.$$

In order to show that $I + \alpha$ is a Pr matrix, consider the principal submatrix t formed by deleting the first and last rows and the first and last columns of

$I + \alpha$. Partition t as $[t_{ij}]$, $i, j = 1, 2$, where the t_{ij} are square matrices of order $2b - 1$. A series of elementary transformations will show that t is non-singular. First, subtract the $(i+1)^{st}$ row of $[t_{21}\ t_{22}]$ from the i^{th} row of $[t_{11}\ t_{12}]$ $(i = 1,2, \ldots, 2b-2)$. The resulting t_{12} is non-singular. Now, add the

$$(i+1)^{st} \text{ column of } \begin{bmatrix} t_{12} \\ t_{22} \end{bmatrix} \text{ to the } i^{th} \text{ column of } \begin{bmatrix} t_{11} \\ t_{21} \end{bmatrix} \quad (i = 1, 2, \ldots, 2b-2).$$

The resulting t_{11} is zero and the resulting t_{21} is a non-singular diagonal matrix. Hence, $|t| \neq 0$, that is, $I + \alpha$ and hence $I + A$, are Pr matrices, which completes the proofs of Theorems 1 and 2.

COROLLARY. *If B is a* c.a. *of a non-singular diagonal matrix d, then $I + B$ is a CPr matrix.*

We have already shown that if B is a c.a. of a non-singular diagonal matrix d and if $\mathbf{X}_1$ is a set of linearly independent rows of $I + B$, then the set $\mathbf{X}^1$ of the corresponding columns is also linearly independent. However, B' is a c.a. of d^{-1} and so linear independence amongst a set of columns of $I + B$ implies linear independence amongst the set of the corresponding rows.

We wish to characterize all of the skew-symmetric matrices q which yield the same c.a. P as the skew-symmetric matrix Q which has just been constructed. Certainly, necessary and sufficient conditions that q also yields P are

$$\text{(i)} \quad (q - Q)(I + P) = 0$$
$$\text{(ii)} \quad |A + q| \neq 0.$$

Theorem 3 will provide a simpler set of conditions.

THEOREM 3. *Let P be a* c.a. *of A, having a parameterization as defined by equation* (1). *Then necessary and sufficient conditions that the skew-symmetric matrix q also yields P are*

$$(9') \qquad \text{(i)} \quad (q - Q)(I + P) = 0,$$
$$(9'') \qquad \text{(ii)} \quad \textit{Rank of } q = \textit{Rank of } Q \qquad (= \textit{Rank of } I - P).$$

Let $P = (A + q)(A - q)$. Then $2q = (A + q)(I - P)$ proving $(9'')$. Furthermore, equation $(9')$ follows immediately from equation $(3')$.

Conversely, let q satisfy $(9')$ and $(9'')$. By $(9')$, q_1 (the analogue of Q_1, formed by applications of Lemma 1 and similar transformations on q) is given by (10) for some X, Y and Z. Let I_s denote the identity matrix of order s (s is the rank of $I_2 + B$). Now, partition B as $[B_{ij}]$, $(i, j = 1,2)$, such that B_{22} is of order s and $I_s + B_{22}$ is non-singular. Define R by $B_{21} = (I_s + B_{22})R$. Then

$$(I_2 + B)[I_s - R']' = 0, \ (I_2 - B)[I_s - R']' = 2[I_s - R']'.$$

Hence, if we set

$$S = \begin{bmatrix} I_2 & 0 & 0 \\ 0 & I_5 & 0 \\ \frac{1}{2} Y \delta^{-1}[I_5 - R'] & 0 & I_4 \end{bmatrix},$$

then

$$Sq_1S' = \begin{bmatrix} (I_2 + B')d(I_2 - B) & -K' & 0 \\ K & X & 0 \\ 0 & 0 & Z \end{bmatrix},$$

which has rank equal to that of Q_1 if and only if $|Z| \neq 0$. However, we have previously shown that $|Z| \neq 0$ if and only if $|A + q| \neq 0$.

By considering matrices whose elements are taken from an arbitrary field of characteristic two, we can exhibit a counterexample to Theorem 1. It is easily seen that the matrix

$$P = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

is a c.a. of the symmetric matrix $A = 1 \dotplus 0$ and that $I + P$ spans the same row space as A. Furthermore, $|P| = +1$. However, for any skew-symmetric matrix Q, $(A + Q)^{-1}(A - Q) = I$.

4. The complex case. Since the proof of the theorem, analogous to Theorem 1, in which the underlying field is the complex field and in which transpose is replaced by conjugate transpose, is slightly simpler but extremely similar to the proof of Theorem 1, we shall only state the theorem and not repeat the proof.

THEOREM 4. *If A is a (not necessarily non-singular) Hermitian matrix and if Q is a skew-Hermitian matrix such that $A + Q$ is non-singular, then equation* (1) *defines a* c.a. *P of A having the property that A and $I + P$ span the same row space.*

Conversely, if P is a c.a. *of A having the property that $I + P$ and A span the same row space, then there is a skew-Hermitian matrix Q such that P is given by equation* (1).

REFERENCES

1. H. Schwerdtfeger. *Introduction to Linear Algebra and the Theory of Matrices* (Groningen, 1950).
2. H. Taber. *On the automorphic linear transformation of an alternate bilinear form.* Math. Ann., *46* (1895), 561–583.
3. J. Williamson. *On the normal forms of linear canonical transformations in dynamics.* Amer. J. Math., *59* (1937), 599–617.
4. ———. *Quasi-unitary matrices,* Duke Math. J., *3* (1937), 715–725.

The University of Wisconsin
and
The University of Rochester

SOME REMARKS CONCERNING CATEGORIES AND SUBSPACES

J. R. ISBELL

Introduction. This paper is primarily a brief elaboration on the axioms for a *bicategory* introduced in **(3)**. From this point of view, the main aim is the development of the structure of certain systems of topological and uniform spaces, and the present paper merely points out some very general properties which follow from axioms so weak that they are satisfied by any system likely to be considered. However, from the point of view of the general theory of categories, the main content of this paper consists of a definition and certain technical observations which tend to justify the particular axioms used. The following remarks must serve as introduction for both viewpoints.

A category is an algebroid system analogous to a group. Rather than define a category here we define a category of mappings, which is analogous to a group of transformations. A category of mappings (Q, A, B) consists of a collection Q of sets, called *spaces*, a collection A of functions on spaces into spaces, called *mappings*, and a subset B of $A \times A \times A$ consisting of those triples (f, g, h) such that h is the composed mapping $g \circ f$. The sole requirements are that A is closed under composition and contains, for each space X in Q, the identity function $i: X \to X$. In particular, a group A of transformations on a set X forms a category if we take $Q = \{X\}$ and $B = \{(a, b, ba) | a \text{ and } b \text{ in } A\}$.

With many mathematical structures there are naturally associated categories of mappings. For example, with a collection Q of groups we may associate the category A of all homomorphisms on elements of Q into elements of Q. Many natural correspondences involve transformations of one category into another. This is most familiar in algebraic topology; for example, in associating with a space X a homology group $H(X)$ one also associates to each continuous function $f: X \to Y$ a homomorphism $f': H(X) \to H(Y)$. However, the phenomenon is also common in general topology. For example, the Stone-Čech compactification induces such a transformation. Passage from a space X to its ring of real-valued continuous functions $C(X)$ is an instance of a *contravariant* transformation, in that a function $f: X \to Y$ induces a homomorphism in the opposite direction, $f^*: C(Y) \to C(X)$.

A category is in the first place an abstract algebra, or at least an abstract structure resembling an algebra. The first section of this paper establishes

Received January 11, 1957. Research supported in part by the Office of Naval Research under Contract N7onr 41904, George Washington University Logistics Research Project, and in part by a National Science Foundation fellowship.

some simple propositions on homomorphisms and congruence relations in categories. For example (as in algebras), a one-to-one homomorphism is an isomorphism (1.1). However, the general homomorphism is not determined by the congruence relation which it induces (1.2).

For the applications one may wish to consider more structure than is given by the law of composition. For example, in a category of mappings (Q, A, B) one may be concerned with those mappings $f: X \to Y$ which embed X as a subspace of Y. Such mappings have special properties expressible in terms of the algebra of composition; for example, such an f always satisfies the cancellation law $fg = fh$ implies $g = h$. In any particular category the concept of "subspace" may or may not be definable in terms of the algebra of composition. The question has been considered how to impose axioms on a subset I of A so that I may reasonably be interpreted to be the set of all embeddings of subspaces. Axioms have been given (involving more than this) by MacLane in **(5)** and by the author in **(3)**; the more elaborate structure so defined is called a *bicategory*.

The second section of this paper is a study of conditions on a subset I of a category A in order that A may be represented as a category of mappings in such a way that I is just the class of embeddings of subspaces. First we consider conditions for an *isomorphic* representation of A so that mappings in I become actual inclusion functions $f: X \to Y$, where X is a subset of Y. Five conditions are taken from **(5)**, and a sixth is shown to be necessary. We sketch a proof that the six conditions are sufficient (2.2). But at this point we have an already formidable battery of axioms, which still do not cover all the primitive terms of bicategory theory. In search of simplicity we turn in another direction.

A *skeleton* of a category is a certain kind of subcategory. Given a category of mappings (Q, A, B), a skeleton is obtained as follows. A mapping $f \in A$ is an *isomorphism* provided f is one-to-one onto and the function f^{-1} is an element of A. Then let K be a subset of Q consisting of just one space from each isomorphism type. The set of all mappings in A whose domain and range are in K is a skeleton of A. (Any two skeletons of A are isomorphic categories.) Then we define two categories to be *coextensive* if they have isomorphic skeletons. We seek conditions on a subset I of A in order that A be coextensive with a category of mappings in such a way that the mappings in I correspond to functions gfh, where f is an inclusion function and g and h are isomorphisms. The necessary and sufficient conditions are (1) for f in I, $fg = fh$ implies $g = h$, and (2) I contains all isomorphisms and is closed under composition with isomorphisms (2.4).

The third section of the paper gives the bicategory axioms of **(3)**, with a few elementary consequences and some discussion of examples. In particular, modulo the identification of coextensive categories, the subspace concept in groups and in compact spaces is definable in terms of the algebra of composition. This is not true in MacLane's more delicate theory **(5)**. It is not

proposed to supplant categorical isomorphism with coextension. However, it is suggested that considerable work remains to be done, at least in general topology, in the study of coextensive invariants of categories of continuous functions. For such work the present system of axioms has substantial advantages.

The author is indebted for discussions and suggestions, particularly to Saunders MacLane, and also to James Case, Pierre Conner, Melvin Henriksen, and Dana Scott.

I. Categories. We begin with a formal definition of a category which is virtually the same as the one given in **(2)**.

Definition. A *category* is an ordered pair (A, B) of sets, where B is a subset of $A \times A \times A$ and the following conditions are met.

(a) For each f, g, in A, there is at most one h in A such that (f, g, h) is in B; such an h is designated gf.

(b) For each f in A there exist (i) at least one i in A such that if exists and for all x in A, (1) if ix exists then $ix = x$, and (2) if xi exists then $xi = x$; and (ii) at least one j in A satisfying (1) and (2) and such that fj exists.

(c) (i) If fg and gh exist then $(fg)h$ exists, $f(gh)$ exists, and $(fg)h = f(gh)$; (ii) if $(fg)h$ exists then gh exists; (iii) if $f(gh)$ exists then fg exists.

Uniqueness of the i and j of condition (b) follows, as shown herewith. Let us call an element of A an *identity* if it satisfies the conditions (1) and (2). Suppose $i'f$ exists, and $if = f$. Then $i'f = i'(if)$, and $i'i$ exists by (c). If i is an identity then $i'i = i'$ and i' is not an identity unless $i' = i$.

The axioms are satisfied by any semigroup A with unit. However, that is not the most interesting sort of category. The sort of "category" one would like to study is illustrated by the "collection" of all continuous functions $f: X \to Y$, where X and Y are compact spaces and the composition gf is the functional composition $g \circ f$. Such a collection of course involves the paradoxes of set theory.

A perfectly proper description of categories which are too large to be sets can be given, for example, in terms of Hilbert-Bernays set theory. Eilenberg and MacLane pointed this out in **(2)**, and MacLane actually carried it out in **(5)**. Until the theory develops further it seems reasonable to duck the complications involved in this development, so far as possible. In this paper we can do this, in spite of the fact that we are concerned primarily with applications to proper classes. All the theorems are stated for sets. In most cases the application may properly be interpreted along the following simple line: a proposition asserted, for example, for (the class of) all continuous functions may as well be asserted for (the set of) all continuous functions on spaces whose points are a subset of a fixed set S, for each S. This interpretation is not right for the representation theorems; generalization of 2.4 or of 3.5, for example, to apply to proper classes, is an unsolved problem. Aside from this, the entire argument could be carried out in Zermelo set theory.

To introduce another convention: a category may reasonably be regarded as an ordered triple (Q, A, B), where Q is a collection of *spaces*, A a collection of *mappings*, and B a subset of $A \times A \times A$ giving the law of composition in A. Since the algebra of mappings is the center of interest, we have defined a category as a pair (A, B); one or another set Q of spaces may be considered to provide a *representation*. In conformity with algebraic (and topological) usage, we may speak of A alone as the category, letting the law of composition be understood. However, in examples, we may name Q alone, as in "the category of all groups"; in such a case it is to be understood that A consists of the usual mappings of such objects (if Q consists of the groups then A consists of their homomorphisms), and B gives the usual law of functional composition. Note, though, that other classes of mappings may be explicitly indicated, and in particular, in speaking of a subcategory there is no presumption that all possible mappings are included. For example, it may be convenient to refer to a subcategory consisting of one group G, one subgroup H, and one isomorphism of H into G.

Yet another convention: a *function* $f: G \to H$ is an ordered triple (f, G, H), where f is a single-valued relation in $G \times H$, G is the set of arguments of f, and H contains the set $f(G)$ of values of f. H is called the *range* of f; $f(G)$ has no particular name. In loose talk we may call $f(G)$ the image of G or of f, but we need the technical term *image* for another use.

In an abstract category A the terms *domain* and *range* are applied to the handiest objects which suggest the domain and range of a function. Specifically, the *domain* of f, $\delta(f)$, is that identity i such that fi exists; and the range $\rho(f)$ is that identity j such that jf exists.

Note that a category may be regarded as an "algebra" with one operation fg, or with three operations, including δ and ρ. In either case it is not precisely an algebra, since fg is not defined for all pairs. However, with the three operations one has a structure which is quite nearly algebraic; the necessary and sufficient condition for the existence of fg is that $\delta(f) = \rho(g)$. (Proof omitted.) One could throw in a zero and define $fg = 0$ if fg is not otherwise defined; however, $\delta(0)$ and $\rho(0)$ would raise new problems. So far as is known, the structure of categories is not adequately described by any strictly algebraic formulation.

Eilenberg and MacLane have shown (**2**, Appendix) that every abstract category may be represented as a category of sets and functions. Specifically, a *concrete category* is defined as an ordered pair (Q, A), where A is a set of functions on elements of Q into elements of Q, and the axioms are

0. For each f in A, the domain and range of f are in Q.

1. Every identity function $i: X \to X$ whose domain is in Q is a member of A.

2. For any $f: X \to Y$ and $g: Y \to Z$ in A, $gf: X \to Z$ is in A.

Either Q or A may be called the category when the meaning is apparent. Every concrete category (Q, A) determines an abstract category (A, B) in the obvious way; the representation theorem is that every abstract category

(A',B') is *isomorphic* with such a category, where isomorphism has the obvious meaning.

Specifically, an isomorphism of (A', B') upon (A, B) consists of a one-to-one correspondence τ of A' onto A such that the induced correspondence of $A' \times A' \times A'$ onto A^3,

$$(f, g, h) \to (\tau(f), \tau(g), \tau(h)),$$

maps B' onto B. A *homomorphism* is a mapping $\tau: A' \to A$ satisfying (a) if fg exists in A' then $\tau(f)\,\tau(g)$ exists and is $\tau(fg)$, and (b) if f is an identity in A' then $\tau(f)$ is an identity in A. One may replace (b) (in the presence of (a); proof omitted) with the conditions $\delta\tau = \tau\delta$ and $\rho\tau = \tau\rho$. A *subcategory* of A is a subset closed under composition, δ, and ρ. Clearly every intersection of subcategories is a subcategory; thus every subset generates a subcategory, and in particular for each homomorphism $\tau: A \to A'$ there is a least subcategory containing $\tau(A)$, which is called the *image* of A under τ. The homomorphism τ also determines an equivalence relation $\mathfrak{r}$ in A, $x\mathfrak{r}y$ if $\tau(x) = \tau(y)$; an equivalence relation obtainable in this way is called a *congruence relation*. A homomorphic image B of A is called an *identification category* of A, and $\tau: A \to B$ an *identification mapping*, in case the following is true: whenever $\sigma: A \to C$ is a homomorphism such that the congruence relation $\mathfrak{s}$ determined by σ contains the congruence relation $\mathfrak{r}$ determined by τ, then there exists a homomorphism $\xi: B \to C$ such that $\xi \circ \tau = \sigma$.

The rest of this section is devoted to establishing the following results.

1.1. (First Isomorphism Theorem). *If $\tau: A \to A'$ is a one-to-one homomorphism then A is isomorphic with its image under τ.*

1.2. *Every homomorphism determines an identification category, not necessarily isomorphic with the image.*

1.3. *The congruence relations on A form a complete lattice $L(A)$. However, if $\mathfrak{r}$ is a particular member of $L(A)$, and A' the identification category determined by $\mathfrak{r}$, the lattice $L(A')$ and the sublattice of $L(A)$ consisting of all relations containing $\mathfrak{r}$ need not be isomorphic.*

1.4. *For homomorphisms*

$$\tau : A \to A', \quad \alpha : A' \to A'', \quad \beta : A' \to A'',$$

if $\alpha \circ \tau = \beta \circ \tau$ then α and β coincide on the image of A. Hence if $\tau: A \to B$ is an identification mapping and $\sigma: A \to C$ is a homomorphism divisible by τ then the solution of $\xi\tau = \sigma$ is unique.

1.5. *Every homomorphic image of a category A is an identification category of an identification category of A.*

Proposition 1.1 is valid for algebras and is sometimes called the First Isomorphism Theorem. Sometimes such names are applied to certain theorems which are significant only for systems having a zero. At any rate, the negative

statements in 1.2 and 1.3 assure that none of the results commonly called the Second Isomorphism Theorem is valid for categories.

Proof of 1.1. Let τ be a one-to-one homomorphism of A onto the subset B of A'. Then B is closed under δ and ρ. (This is true even if τ is not one-to-one.) In B, the general element has the form $\tau(x)$; and

$$\tau(x)\ \tau(y) \text{ exists in } A' \Leftrightarrow \delta\tau(x) = \rho\tau(y) \Leftrightarrow \tau\delta(x) = \tau\rho(y)$$
$$\Leftrightarrow \delta(x) = \rho(y) \Leftrightarrow xy \text{ exists in } A \Leftrightarrow \tau(x)\ \tau(y) = \tau(xy) \text{ in } B.$$

Therefore B is a subcategory, τ is one-to-one onto B, and $\tau\colon A \to B$ is an isomorphism.

For 1.2 and 1.3 we need the lemma

1.6. *An equivalence relation in a category A which determines the set C of equivalence classes c is a congruence relation if and only if*

(1) *the set product cd of any two members of C is a subset of a member of C, and*

(2) *the sets $\delta(c)$ and $\rho(c)$ are subsets of members of C.*

Proof. Clearly a congruence relation has these properties. Conversely let the partition C satisfy (1) and (2). Let the category B consist of the members of C, and other elements to be described. For $c \in C$, $\delta(c)$ in A is a subset of some element of C; on the other hand, the set $\delta(c)$ is not empty, and thus it lies in a unique member $\delta'(c)$ of C. Similarly ρ in A induces an operator ρ' in C. Altogether let B consist of all ordered n-tuples (words) of elements of C, $(c_1, \ldots, c_n)$, such that for $1 \leqslant i \leqslant n-1$, $\delta'(c_i) = \rho'(c_{i+1})$, but the product in A of the sets c_i, c_{i+1}, is empty. (That is, $\delta(c_i)$ and $\rho(c_{i+1})$ are disjoint subsets of the same element of C.) Define

$$\delta'(c_1, \ldots, c_n) = \delta'(c_n), \quad \rho'(c_1, \ldots, c_n) = \rho'(c_1).$$

The product in B of $(c_1, \ldots, c_n)$ and $(b_1, \ldots, b_m)$ is defined if and only if $\delta'(c_n) = \rho'(b_1)$. If $\delta(c_n) \cap \rho(b_1) = 0$ then the product is $(c_1, \ldots, c_n, b_1, \ldots, b_m)$. Otherwise $c_n b_1$ is a non-empty subset of a unique member d of C; and, suppressing an induction, we describe the product as the word $(c_1, \ldots, c_{n-1}, d, b_2, \ldots, b_m)$, contracted as far as possible by further multiplication.

It is easily seen that we have defined a category B. The function $\tau\colon A \to B$ which takes each member of A to its C-equivalence class is a homomorphism, and thus C defines a congruence relation. Furthermore, B is an identification category. We have finished the proof of 1.6 and begin on the

Proof of 1.2. Given the situation above, with the homomorphism $\tau\colon A \to B$; and given a homomorphism $\sigma\colon A \to D$ constant on each equivalence class c of the partition C; to construct a homomorphism $\xi\colon B \to D$ such that $\xi\tau = \sigma$. For one-letter words $c \in B$, let $\xi(c)$ be the constant value of $\sigma(x)$, for any $x \in c$ in A. Since σ is a homomorphism, therefore

$$\delta\xi(c) = \xi\delta(c), \qquad \rho\xi(c) = \xi\rho(c).$$

Then if (c_1,c_2) is a word in B, necessarily $\xi(c_1)\xi(c_2)$ exists in D. Define $\xi(c_1, c_2)$ to be $\xi(c_1)\xi(c_2)$; and so on by induction. By definition $\xi\tau = \sigma$, and clearly ξ is a homomorphism.

That the identification category need not be isomorphic with the image is perhaps obvious, but we give an example. The homomorphism cannot be one-to-one on identities, for the freedom in the image arises only where new products are defined. Accordingly consider the category A with four elements, x_0, x_1, y_0, y_1;

$$\rho(x_i) = \delta(x_i) = x_0, \qquad \rho(y_i) = \delta(y_i) = y_0, \qquad i = 0, 1;$$

thus x_0 and y_0 act as identities; and finally, $x_1^2 = x_1$, $y_1^2 = y_1$. (A typical realization of A is on a pair of linear spaces, each with its identity mapping and one projection upon a proper subspace.) Consider the category B with four elements, z_0, z_1, z_2, z_{12}, all idempotent, z_0 an identity, z_{12} a zero, $z_1z_2 = z_2z_1 = z_{12}$. There is a homomorphism $\tau: A \to B$ given by

$$\tau(x_0) = \tau(y_0) = z_0, \qquad \tau(x_1) = z_1, \qquad \tau(y_1) = z_2.$$

B is the image; but the identification category determined by τ is neither finite nor commutative.

For the proof of 1.3, it is clear from 1.6 that every intersection of congruence relations is a congruence relation. Hence any equivalence relation generates a least containing congruence relation, and $L(A)$ is a complete lattice. In the example in the proof of 1.2, $L(A)$ is a finite lattice, while the identification category clearly has infinitely many congruence relations. This finishes 1.3.

The proof of 1.4 is a trivial induction.

For 1.5 we establish a lemma.

1.7. *Let $\tau: A \to B$ be a homomorphism which is one-to-one on identities. Then the identification category determined by τ is isomorphic with the image of A under τ.*

Proof. From the proof of 1.6 we see that the homomorphism τ^* of A upon the identification category A' is onto unless for some equivalence classes, c_1, c_2, the sets $\delta(c_1)$ and $\rho(c_2)$ are disjoint subsets of the same equivalence class. This is impossible when τ is one-to-one on identities. But then the quotient homomorphism $\xi: A' \to B$ is one-to-one, since otherwise $\tau = \xi\tau^*$ would determine a larger congruence relation on A. Then 1.1 applies, and 1.7 is proved.

Proof of 1.5. From the proof of 1.6 we see that every identity in an identification category is the image of an identity in A. Hence the induced mapping of the identification category upon the image is one-to-one on identities, 1.7 applies, and 1.5 is proved.

II. Subspaces. In **(3)** there is given a simplified version of MacLane's axioms for a *bicategory* (crudely: a category with subspaces), which will be used in the concluding portion of this paper. The simplified version has a

somewhat different motivation than (may be presumed for) the original, and it seems likely that a combination of the two may survive. Basically, the simplification involves a broader notion of equivalence. In **(5)** MacLane investigates properties invariant under *isomorphism*. Below we define a relation of *coextension*, and we shall be concerned with coextensive invariants. Isomorphic categories or bicategories are coextensive, but not conversely.

This section illustrates the two viewpoints—mainly the cruder one—in examining the question what axioms must be imposed on subspaces in order that they behave like subsets.

A function $f: X \to Y$ is called an *inclusion* function provided X is a subset of Y and $f(x) = x$ for all x in X. Given a category A and a subset I of A, under what conditions can A be represented as a concrete category so that the mappings in I, and no others, become inclusion functions? Five clearly necessary conditions are

(1) every identity is in I,

(2) I is closed under composition (by (1) and (2), I is a subcategory),

(3) $f = gh$ with f and g in I implies h is in I,

(4) $fg = fh$ with f in I implies $g = h$, and

(5) I contains at most one element with given domain and range.

These conditions have been recognized by MacLane and incorporated mutatis mutandis into his axioms **(5)**. A sixth condition is necessary and, for the immediate question, sufficient; but we shall merely sketch the proof (2.2).

Let two elements of A, f and g, be called *equivalent* if there is a finite chain $(h_1, \ldots, h_n)$, $h_1 = f$, $h_n = g$, such that for $1 \leqslant i \leqslant n - 1$, either $h_i = j_i h_{i+1}$ or $h_{i+1} = j_i h_i$, for some j_i in I. Supposing I to be the set of inclusion functions of a concrete category, equivalence of f and g implies that f and g have the same domain and values. Therefore we may demand (6) two equivalent elements of A having the same range are identical.

Remark 2.1. The conditions (1)—(5) do not imply (6). In fact, one can construct a system satisfying all the axioms and conventions of **(5)** for bicategories, in which the class of *injections* (in the language of **(5)**) does not satisfy (6). The construction is straightforward but too tedious and unsurprising to give here.

Remark 2.2. The conditions (1)—(6) imply that A is isomorphic with a concrete category in such a way that I corresponds precisely to the inclusion functions. The reasons for omitting the somewhat lengthy proof are (a) that the result seems to be useless both in the context of MacLane's theory, where it is not strong enough, and in the context of this paper, where it is irrelevant; and (b) it is a mere modification of the Eilenberg-MacLane representation of **(2)**. In fact, A is partitioned into equivalence classes by the relation of equivalence defined above; carry through the Eilenberg-MacLane construction and then choose a representative f_0 of each equivalence class $[f]$, and replace each occurrence of f by f_0. The representation is preserved

because of assumption (6), the elements of I become inclusion functions because of (4), and there are no other inclusion functions because of (1)—(3). (Condition (5) is an easy consequence of (6).)

In any category A, f is said to be an *isomorphism* if there exists a mapping f^{-1} in A such that ff^{-1} is an identity and $f^{-1}f$ is an identity. In a concrete category we call f an *injection* if f has the form gih, where i is an inclusion function and g and h are isomorphisms. Two injections f and g are *equivalent* if there is an isomorphism j such that $fj = g$. An equivalence class of injections into X is called a *subspace* of X. (Thus a subspace has a fixed range, and (speaking imprecisely) a fixed "image," but the domain is determined only up to isomorphism.)

One might ask under what conditions a category A can be represented with a prescribed family I of injections. Clearly I must contain all isomorphisms and be closed under composition with isomorphisms. Further, the cancellation condition (4) above must hold. More arcane properties can be found, for example, if X has $\mathbf{m}$ subspaces isomorphic with Y ($\mathbf{m}$ a cardinal number), then A contains $\mathbf{m}$ spaces isomorphic with Y. But this is not what we want.

Accordingly we define a *skeleton* K of a category A as follows. A subcategory S of A is *full* in case the hypotheses $\delta(f) \in S$ and $\rho(f) \in S$ imply $f \in S$. Two identities, i, j, are *isomorphic* or *equivalent* if there exist isomorphisms f, g, such that $fg = i$ and $gf = j$. Then a *skeleton* is a full subcategory K including exactly one identity from each equivalence class.

2.3. *All skeletons of a category A are pairwise isomorphic.*

Proof. Let K, K' be two skeletons of A. For each identity i in K there is exactly one equivalent identity i' in K', and at least one isomorphism f in A such that $f^{-1}f = i, ff^{-1} = i'$. For each i in K choose one such f. For any g in K, let f_1 be the isomorphism associated with $\delta(g)$, f_2 the isomorphism associated with $\rho(g)$. Then $g' = f_2gf_1^{-1}$ is in K', and the transformation $g \to g'$ is evidently an isomorphism.

We define two categories to be *coextensive* if they have isomorphic skeletons. In the ordinary parlance of algebra and topology, outside of homology theory, the distinction between coextensive categories is commonly ignored. This is not to say that it ought to be ignored; but one may properly investigate those properties of categories which are coextensive invariants.

If we prescribe a class I of *injections* in a category A, and I satisfies the modest requirement of including all isomorphisms and being closed under composition with isomorphisms, then for any skeleton K of A the family $K \cap I$ also has these properties. Further, I is determined by $K \cap I$. (Proof omitted.) Note, though, that the definition of injections for a concrete category does not relativize to a skeleton in general. If we agree that in a skeleton of a concrete category (Q, B) the term *injection* is to be defined by reference to the whole category, then we have

2.4. *A category A with a distinguished subset I is coextensive with a concrete category under an isomorphism of skeletons identifying I with the class of injections, if and only if*

(1) *I contains all isomorphisms and is closed under composition with isomorphisms, and*

(2) *for f in I and g and h in A, $fg = fh$ implies $g = h$.*

We preface the proof with some remarks and a lemma. The construction is a modification of that of Eilenberg-MacLane **(2)**; like that one, it does not extend to proper classes. For the lemma, let us introduce the term *projection* in a somewhat unusual way. Relative to a prescribed class I of injections, f is a *projection* if the hypothesis $f = gh$, where g is an injection, implies g is an isomorphism.

2.5. *Under the conditions of* 2.4, *every isomorphism is a projection.*

Proof. It suffices to consider identities, for if k is an isomorphism having a factorization fg, f a proper injection, then the range of k is $f(gk^{-1})$. If the identity i is fg, f in I, we consider the mapping gf and its domain (= its range) j. We have $f(gf) = if = f = fj$; hence by condition (2), $gf = j$. Thus f must be an isomorphism, as was to be shown.

Proof of 2.4. The necessity of conditions (1) and (2) is clear. For the converse, choose a skeleton K of A, and let $J = I \cap K$. Let Q be the set of all identities of K. We must construct a concrete category (S, M) of spaces and mappings, having a skeleton K' involving a set Q' of spaces, with an isomorphism of K upon K' identifying J with the injections.

For each i in Q, let X_i be the set of all mappings in K with range i. Let $Q' = \{X_i | i \in Q\}$. For each f in K, define $f' : X_{\delta(f)} \to X_{\rho(f)}$ by $f'(g) = fg$. Let $K' = \{f' | f \in K\}$. For each element f of J, f' is one-to-one, by condition (2); let the set $f'(X_{\delta(f)})$ be an element of S, and let S consist precisely of all these sets (f ranging over J) and all the elements of Q'. Let M consist of

(a) the functions in K',

(b) for each element f of J, the function f^* agreeing with f' on the domain $X_{\delta(f)}$, with range (cut down to) $f'(X_{\delta(f)})$, the function $(f^*)^{-1}$, and the inclusion function $i : f'(X_{\delta(f)}) \to X_{\rho(f)}$; and finally

(c) finite compositions of functions in (a) and (b).

Let us call the functions under (b) *b-mappings* (b for basic).

(S, M) satisfies the axioms **0**, **1**, **2**, for a concrete category, obviously. (The identities on spaces not in Q' are compositions $f^*(f^*)^{-1}$.) Next, if f is in J but not an isomorphism then the function f' is not onto; in fact, the identity $\rho(f)$ is not in the range of f^*, by 2.5. This shows that each b-mapping either is in K' or has a domain or range not in Q'. By induction, every mapping $g \in M$ whose domain and range are in Q' is an element of K'. Clearly each space in S is isomorphic with at least one space in Q'; K includes no two isomorphic identities; if f is not an identity in K then $f\delta(f) \neq \delta(f)$, and therefore

K' includes no isomorphism between two different spaces. Therefore K' is a skeleton of M.

Finally, it is clear that K and K' are isomorphic and every element f of J corresponds to an injection $f' = if^*$. An induction shows that every inclusion function in M is a b-mapping; and then every injection gih in K' corresponds to an element of J. This completes the proof of 2.4.

III. Bicategories. In a category A, a mapping f is *left cancellable* if $fg = fh$ implies $g = h$, *right cancellable* if $gf = hf$ implies $g = h$. A *bicategory* $\mathscr{C}$ is an ordered triple (A, I, P), where A is a category and I and P are subsets of A whose members are called *injections* and *projections*, respectively, and

3. Both I and P are subcategories containing all the isomorphisms.

4. Every mapping f in A is a composition gh of a projection h and an injection g. This decomposition, or factorization, is essentially unique; that is, the only other such expressions of f are of the form $(gj^{-1})(jh)$, where j is an isomorphism.

5. (a) Every injection is left cancellable. (b) Every projection is right cancellable.

The axioms are stated for an abstract category, but clearly

3.1. *Every bicategory (A, I, P) is coextensive with a concrete category under a correspondence representing I as the set of injections.*

For the axioms imply the hypothesis of 2.4. As for P, **(3**, Lemma 2.0**)** states

2.0. *In any bicategory, the mapping f is a projection if and only if $f = gh$, with g an injection, implies g is an isomorphism.*

One can also derive some of the results of **(5)** from these axioms. In particular, if $f = gh$ with f and g in I then h is in I, because of the uniqueness clause in Axiom 4. Thus we have the first four of the six properties of inclusion functions listed at the beginning of the previous section. Since the axioms **3**—**5** are preserved under passage to a skeleton K of A (replacing I with $I \cap K$, P with $P \cap K$), but the fifth and sixth properties in the cited list are not so preserved, therefore the axioms imply all those properties of inclusion functions which are coextensive invariants. I do not know if this remark can be made precise. (It cannot be done by constructing a coextensive concrete bicategory with I = inclusions, because there could be no non-identical isomorphisms.) But perhaps it conveys the idea.

What sort of category A can be made into a bicategory (A, I, P)? To begin with, it suffices if A is a concrete category in which for every mapping $f: X \to Y$ the set $f(X)$ is a space; the definitions are obvious and the proof is omitted. Any category coextensive with a bicategory is, in a natural way, a bicategory. Let us consider a more restrictive condition.

6. Every mapping which is left and right cancellable is an isomorphism.

3.2. *A category satisfying Axiom* **6** *can be made into a bicategory in at most one way.*

Proof. Suppose A satisfies **6**, and (A, I, P) and (A, I', P') are two bicategories. For any f in I, consider the factorization $f = gh$, g in I', h in P'. Since f is left cancellable, so is h (compute); since h is in P', it is also right cancellable and hence an isomorphism. Thus I is a subset of I'; by the same argument, I' is a subset of I, and by 2.0, $P = P'$ as well.

The one way is, of course, with I = all left cancellable mappings, P = all right cancellable mappings. The conditions for this to make a bicategory are Axioms **3**, **4**, and **5**; but **3** and **5** are trivial here. Axiom **4**, in this case, implies **6**; hence we may reduce the axioms to

3.3. The necessary and sufficient conditions on a category A in order that (A, I, P) form a bicategory, where I consists of all left cancellable mappings and P of all right cancellable mappings, are (1) I and P generate A, (2) for i in I and p in P such that pi exists, there are i' in I and p' in P such that $i'p' = pi$, and (3) for i and i' in I, if i' is not ij for any isomorphism j, then there do not exist p and p' in P such that $ip = i'p'$. If I' and P' are subcategories of I and of P, respectively, each containing all isomorphisms, then conditions (1)—(3) applied to I' and P' are necessary and sufficient in order that (A, I', P') be a bicategory.

The proof is omitted.

Axiom **6** may be regarded as a form of the First Isomorphism Theorem. It holds in many interesting categories; for example, in any *exact* category in the sense of **(1)**, and in any equationally definable class of algebras with zero, in the sense of **(4)**, where of course the mappings are the homomorphisms. In each case the proof is a routine exercise in the relevant theory. That the axiom is invalid for the most general types of algebras, and for some other types of systems, is illustrated by a rather trivial example. Consider all those algebras, (S, O), where S is a ground set and O a set of finitary operations on S, which is empty, and the (non-existent) operations in O are subjected to the one requirement $x = y$. There are precisely two algebras and three homomorphisms in any skeleton of this category, and Axiom **6** is clearly false. Thus if the axiom is to be satisfied one must exclude this sort of pathology. However, a category may contain this example and still satisfy Axiom **6**, as is shown by the compact Hausdorff spaces.

Another illustration is given by the category of all categories; precisely, the proper class $A \cup B$, where A is the class of all homomorphisms $f: X \to Y$, X and Y being categories which are sets, and B is the obvious subclass of $A \times A \times A$. In proving this, let us designate homomorphisms of categories by Latin letters, elements by Greek letters. Suppose $f: X \to Y$ is cancellable on both sides. Then f is one-to-one on identities; for if $f(\alpha) = f(\beta)$, α and β identities, then α and β form a two-element category Z which is mapped into X by the inclusion function $i: Z \to X$, and another homomorphism $j: Z \to X$ is defined

by $j(\alpha) = j(\beta) = \alpha$. Here $fi = fj$ but $i \neq j$, a contradiction. Therefore if $f(\alpha) = f(\beta)$ for any two elements α and β of X, we may conclude $\delta(\alpha) = \delta(\beta)$ and $\rho(\alpha) = \rho(\beta)$. If $\delta(\alpha) \neq \rho(\alpha)$ then there is a four-element category consisting of α, β, $\delta(\alpha)$ and $\rho(\alpha)$, which clearly has two different homomorphisms into X which have the same composition with f. There remains the case $\delta(\alpha) = \rho(\alpha) = \gamma$. Then α, β, and γ generate a semigroup Z with unit γ. Let W be the free semigroup with unit on two generators, σ, τ. Then W is a category, and there exist two homomorphisms $h\colon W \to Z$, $k\colon W \to Z$, determined by the conditions $h(1) = k(1) = \gamma$, $h(\sigma) = k(\tau) = \alpha$, $h(\tau) = k(\sigma) = \beta$. Composing h and k with the injection $i\colon Z \to X$, we obtain two category homomorphisms $ih\colon W \to X$ and $ik\colon W \to X$ such that $ih \neq ik$ but $fih = fik$. The contradiction establishes that f must be one-to-one. Then by 1.1, f is an isomorphism of X upon its image $Z \subset Y$. It remains to show that if Z is a proper subcategory of Y, then there exist a category U and two different homomorphisms of Y into U which coincide on Z. We omit the details of the argument, which turns on constructing a free sum of two copies of Y modulo the identification of the two copies of Z.

Thus the neat structure described by Axioms **0**—**6** is not uncommon. We do not have it, however, in non-compact topological spaces. As noted in **(3)**, what one typically finds in this example (say, all continuous mappings between Hausdorff spaces) is that the category A can be made a bicategory in two ways; once with all left cancellable mappings taken for injections, and again with every right cancellable mapping a projection. The common part, the two-sided cancellable mappings, consists of those one-to-one continuous functions whose image is a dense subspace of the range. The smaller classes of injections and of projections are then respectively the injections (in the ordinary sense) of closed subspaces, and the identification or quotient mappings.

We have avoided the term "quotient." The difficulty is in distinguishing between quotient and image. Now in groups, and in many other examples, the quotient and image in the usual sense are isomorphic; the distinction is a rather delicate one to make in an abstract setting, and the present bicategory axioms cannot do it. For work involving such distinctions one must use the original formulation of MacLane **(5**, see §11**)**. In topology, however, the quotient and image are typically quite different. They arise not in the factorization belonging to one bicategorical structure, but in two different ones. Note that a topological quotient mapping is categorically definable; $f\colon X \to Y$ is a quotient mapping if and only if the equation $f = gh$, with g left cancellable, implies g is an isomorphism. From each mapping h, of course, one can factor out the unique quotient mapping k such that $h = jk$ with j left cancellable. A similar, but more complicated, description of images can be given by reference to the one-point space.

Thus we have discriminated the two main uses of the terms. Clearly they conflict, and we cannot anticipate a revision either in topology or in algebra. We need a term for the blurred quotient-or-image given by projections accord-

ing to Axioms **0**—**5**. Let it be *quotient*; precisely, a quotient is an equivalence class of projections under the equivalence relation defined by $f \sim g$ when $f = ig$ for some isomorphism i. (This is perfectly analogous to the definition of a subspace.)

This choice frees the term "image," which happens to be wanted on several other counts. Some of these are (1) the use in the refined theory of bicategories, (2) the use, at least informally, for sets of values $f(X)$, (3) the use in connection with category homomorphisms (definition preceding 1.1), and (4) the following use. If $\phi = [f]$ is an equivalence class of injections into X, and $g: X \to Y$ a mapping, then the projection-injection factorization of gf yields a subspace of Y which is most naturally called the *image* of ϕ under g.

Now consider the propositions 1.3, on congruence relations, and 1.5, on identification categories. They are partially misleading, considered alone. But now we see that the trouble is that the congruence relations and identification categories have less to do with the categorical structure in this example than in either algebra or topology. If we replaced the concept of an identification category with the concept of a quotient, defined as an image under a mapping having no proper left cancellable left factor, then we should find 1.5 replaced by the proposition "Every homomorphic image is a quotient." Similarly the lattice isomorphism denied in 1.3 could be rediscovered by looking at the lattice of images instead of the lattice of congruence relations.

Next, the definition of a quotient in a bicategory (two paragraphs back) is more than merely analogous to the definition of a subspace; it is dual. The *dual* A^* of any category A is (a category) in one-to-one correspondence with A, $f \leftrightarrow f^*$, such that g^*f^* is defined if and only if fg is defined, and in that case $g^*f^* = (fg)^*$. It follows **(2)** that A^* is a category, $\delta(f^*) = \rho(f)^*$, and $\rho(f^*) = \delta(f)^*$. If (A, I, P) is a bicategory, then (A^*, I^*, P^*) is a bicategory **(3)**, where I^* is the image of P under $f \to f^*$, and P^* is the image of I. That is,

3.4. *Every bicategory has a dual, unique up to isomorphism, which is a bicategory.*

The proof is omitted.

We conclude with an important definition and a sketch of an embedding theorem. A subcategory $\mathscr{D}$ of a bicategory $\mathscr{C}$ is said to be *regular* if it is closed under factorization, i.e. if f is an injection in $\mathscr{C}$ and g a projection in $\mathscr{C}$, and fg is in $\mathscr{D}$, then f and g are in $\mathscr{D}$. A regular subcategory of a bicategory is of course a bicategory with the relativized sets of injections and projections. Every intersection of regular subcategories is regular, and therefore every subcategory (for that matter, every subset) is contained in a least regular subcategory.

3.5. Every concrete category which is a bicategory may be embedded as a regular subcategory of a bicategory satisfying Axiom **6**.

The embedding is an isomorphism; if the concreteness hypothesis is removed, one gets coextension from 2.4. The proof is too long to give here,

mainly because of the first stage. In outline, the first stage is to enlarge the spaces suitably so that mappings which are not injections cease to be one-to-one. The third stage is to introduce a one-point space mapping into every space so that mappings which are not one-to-one cease to be left cancellable; one must precede this by a stage assuring that no existing one-point spaces are confused, which can be done by adding two zeros to each space. For the final stage, consider all pairs (X, Y), X a subspace of Y. In each case form a space Σ consisting of the sum of three copies of Y with the three copies of X identified. Two copies would be enough so that none of the old mappings, not a projection, remains right cancellable; to assure that $Y \to \Sigma$ (each of the three natural mappings) is not right cancellable, provide Σ with a group of six motions permuting the copies of Y. Only six mappings with domain Σ are admitted.

For the first stage, consider the general space X. Let $S(X)$ be the set of all ordered pairs (σ, τ), σ a subspace of X, i.e. an equivalence class of injections $f\colon Y \to X$, and τ an equivalence class of projections $g\colon Y \to Z$ under the relation $g \sim g'$ if $g' = agb$, a and b isomorphisms. The idea is that a mapping $h\colon X \to W$ which is not an injection has a right factor which is a proper projection; something which is surely narrowed by the mapping is the possibility of forming further projections. Thus we should like to transform quotients of X to quotients of W, which we could do directly (for projections h) if the quotients of a given domain formed a complete lattice. As it is, we must build a complete lattice. Accordingly call a subset T of $S(X)$ *residual* provided for each (σ, τ) in T, $f \in \sigma$, $g \in \tau$, T contains the equivalence classes of (1) all pairs (fi, g'), i an injection, g' the projection having the same domain as i arising in factorization of gi, and (2) all pairs (f, hg), h a projection. Replace X with the set X' consisting of the points of X and the residual subsets of $S(X)$. For any mapping $h\colon X \to W$, extend h over X' by taking for $h(T)$ the least residual set in $S(W)$ containing the equivalence classes of all (f', g') such that for some (f, g), $f \in \sigma$, $g \in \tau$, $(\sigma, \tau) \in T$, the following is true. The mapping hf has a factorization $f'k$, k a projection; i.e. σ' is the image of σ. And $g'k = g$, i.e. g' induces g. The empty set is a residual subset of $S(W)$ which may have to be used; however, the padded category is well defined and the straightforward verification of its properties may be omitted.

References

1. D. Buchsbaum, *Exact categories and duality*, Trans. Amer. Math. Soc., *80* (1955), 1–34.
2. S. Eilenberg and S. MacLane, *General theory of natural equivalences*, Trans. Amer. Math. Soc., *58* (1945), 231–294.
3. J. Isbell, *Algebras of uniformly continuous functions*, submitted to Annals of Math.
4. B. Jònsson and A. Tarski, *Direct Decompositions of Finite Algebraic Systems* (Notre Dame, 1947).
5. S. MacLane, *Duality for groups*, Bull. Amer. Math. Soc., *56* (1950), 485–516.

Institute for Advanced Study

COMPLETENESS IN SEMI-LATTICES

L. E. WARD, Jr.

1. Introduction. Let $(X, \leqslant)$ be a partially ordered set, that is, X is a set and $\leqslant$ is a reflexive, anti-symmetric, transitive, binary relation on X. We write

$$M(x) = \{a : x \leqslant a\}, \quad L(x) = \{a : a \leqslant x\},$$

for each $x \in X$. If, moreover,

$$x \wedge y = \sup L(x) \cap L(y)$$

exists for each x and y in X, then $(X, \leqslant)$ is said to be a *semi-lattice*. If $(X, \leqslant)$ and $(X, \geqslant)$ are semi-lattices, then $(X, \leqslant)$ is a *lattice*.

The lattice $(X, \leqslant)$ is *complete* if, for each non-empty subset A of X, elements

$$\wedge A = \sup \bigcap \{L(a) : a \in A\}, \tag{1}$$

$$\vee A = \inf \bigcap \{M(a) : a \in A\} \tag{2}$$

exist. Lattice-completeness has been characterized in various ways; in particular Frink **(4)** showed it equivalent to compactness relative to a natural sort of topology, and Anne C. Davis **(3)** proved it equivalent to an agreeable fixed point condition.

Let us say that a semi-lattice $(X, \leqslant)$ is *complete* provided (1) exists for each non-empty subset A of X. To avoid ambiguity, we shall refer to a structure $(X, \leqslant)$ as being *lattice-complete* or *semi-lattice-complete* whenever it is not clear from context whether $(X, \leqslant)$ is to be regarded as a lattice or a semi-lattice. In what follows, semi-lattice analogues of theorems on lattices due to Frink **(4)**, Tarski **(5)**, and Davis **(3)**, are proved.

2. Topology in partially ordered sets; Frink's theorem. Let $(X, \leqslant)$ be a partially ordered set. The *interval topology* **(2**, p. 60**)** is that topology generated by taking all of the sets $L(x)$ and $M(x)$, $x \in X$, as a subbasis for the closed sets. An element of X is *maximal* (*minimal*) if it has no proper successor (predecessor). A *zero* (unit) of X is an element which precedes (succeeds) all other elements of X.

LEMMA 1. *Let A be a non-empty subset of X, where $(X, \leqslant)$ is a semi-lattice. If $L(a)$ is compact in the interval topology, for some $a \in A$, then the set*

$$L = \bigcap \{L(a) : a \in A\}$$

has a unit.

Received February 19, 1957. Presented to the American Mathematical Society April 20, 1957.

Proof. From **(6**, Theorem 1**)** and the semi-lattice ordering of X, it follows that X has a zero and hence that L is not empty. Again by **(6**, Theorem 1**)** L has a maximal element, x_1. If there exists $x \in L - L(x_1)$ then it may be shown that $M(x) \cap M(x_1)$ is a semi-lattice containing A, and consequently that $M(x) \cap M(x_1)$ has a zero, x_0. It follows that $x_1 < x_0 \leqslant a$ for all $a \in A$, contradicting the maximality of x_1 in L. Therefore, $L \subset L(x_1)$, which is to say that x_1 is a unit for L.

THEOREM 1. *For the semi-lattice $(X, \leqslant)$ to be complete it is necessary and sufficient that, for each $x \in X$, $L(x)$ be compact in the interval topology.*

Proof. Suppose $(X, \leqslant)$ is complete. In view of Alexander's lemma **(1)** it suffices, in order to show $L(x)$ compact, to prove that if $\{x_\alpha : \alpha \in A\}$ and $\{x_\beta : \beta \in B\}$ are subsets of $L(x)$ such that

$$\mathfrak{F} = \{M(x_\alpha) : \alpha \in A\} \cup \{L(x_\beta) : \beta \in B\}$$

is a non-empty collection with finite intersection property, then $\mathfrak{F}$ has a non-empty intersection. We consider two alternatives: either B is empty or it is not. If B is empty, then $x \in \bigcap \mathfrak{F}$; if B is not empty, then by the finite intersection property, $x_\alpha \leqslant x_\beta$ for each $\alpha \in A$ and $\beta \in B$. Therefore, since X is complete,

$$x_\alpha \leqslant \wedge \{x_\beta : \beta \in B\} = x_0$$

for each $\alpha \in A$. Clearly, $x_0 \in \bigcap \mathfrak{F}$.

Conversely, suppose that $L(x)$ is compact for each $x \in X$ and that A is a non-empty subset of X. By Lemma 1,

$$L = \bigcap \{L(a) : a \in A\}$$

has a unit, x_1, and it is clear that $x_1 = \wedge A$.

COROLLARY 1.1 (Frink). *For the lattice $(X, \leqslant)$ to be complete it is necessary and sufficient that X be compact in the interval topology.*

Proof. If $(X, \leqslant)$ is complete as a lattice, then both $(X, \leqslant)$ and $(X, \geqslant)$ are complete as semi-lattices. Therefore, X has a unit, x_1, and $L(x_1) = X$ is compact, by Theorem 1. Conversely, the compactness of X implies the completeness of $(X, \leqslant)$ and $(X, \geqslant)$ as semi-lattices, which is equivalent to the lattice-completeness of $(X, \leqslant)$.

COROLLARY 1.2. *For the semi-lattice $(X, \leqslant)$ to be complete it is necessary and sufficient that $(L(x), \leqslant)$ be a complete lattice, for each $x \in X$.*

Proof. The sufficiency is immediate from Corollary 1.1 and Theorem 1. To prove the necessity, let $x \in X$ where $(X, \leqslant)$ is a complete semi-lattice. By Theorem 1, $L(x)$ is compact. If a and b are elements of $L(x)$, then (see the argument of Lemma 1) $M(a) \cap M(b)$ has a zero, and that zero is $a \vee b$. Thus, $(L(x), \leqslant)$ is a compact (and hence complete) lattice.

3. A theorem of Tarski. If A and B are partially ordered sets, a function $f:A \to B$ is *isotone* if $a_1 \leqslant a_2$ implies $f(a_1) \leqslant f(a_2)$. A *chain* of a partially ordered set is a simply ordered subset. A chain is *maximal* if it is properly contained in no other chain.

The following theorem is due to Tarski **(5)**.

THEOREM T. *Let $(X, \leqslant)$ be a complete lattice. If $f:X \to X$ is an isotone function then the set P of fixed points of f is non-empty; further, $(P, \leqslant)$ is a complete lattice.*

Theorem T fails if the word "lattice" is everywhere replaced by "semi-lattice" (see §4). However, we have

THEOREM 2. *Let $(X, \leqslant)$ be a semi-lattice and let $f:X \to X$ be an isotone function. If X is compact in the interval topology, then the set P of fixed points of f is non-empty. If X is a complete semi-lattice and P is non-empty, then $(P, \leqslant)$ is a complete semi-lattice.*

Proof. If X is compact, it has a zero which precedes its f-image; thus, the set

$$U = \{x : x \leqslant f(x)\}$$

is not empty and contains a maximal chain, C. By the compactness of X, C has a least upper bound u. Since f is isotone, we have $x \leqslant f(x) \leqslant f(u)$ for all $x \in C$, and therefore

$$u \leqslant f(u) \leqslant f(f(u)) \leqslant \cdots.$$

If $u \neq f(u)$ then the maximality of C is contradicted, so that P is non-empty. Now if X is complete as a semi-lattice (and not necessarily compact) and P is non-empty, then by Corollary 1.2, $(L(p), \leqslant)$ is a complete lattice for each $p \in P$. Readily $f(L(p)) \subset L(p)$, so that Theorem T implies that $(P \cap L(p), \leqslant)$ is a complete lattice. By Corollary 1.2 the theorem follows at once.

4. A theorem of Davis. Recently **(3)** Anne C. Davis proved

THEOREM D. *For a lattice $(X, \leqslant)$ to be complete it is necessary and sufficient that every isotone function $f:X \to X$ have a fixed point.*

There exist complete semi-lattices which do not have the fixed point property for isotone functions. The interval $0 \leqslant t < 1$ of real numbers is a simple example. The natural semi-lattice analogue to Theorem D is

THEOREM 3. *For a semi-lattice $(X, \leqslant)$ to be compact in its interval topology it is necessary and sufficient that every isotone function $f:X \to X$ have a fixed point.*

LEMMA 2. *If $(X, \leqslant)$ is a semi-lattice and if every isotone function $f:X \to X$ has a fixed point, then, for each $x \in X$, $(L(x), \leqslant)$ is a lattice.*

Proof. If not, there are elements a, b, and x of X such that a and b precede x and $M(a) \cap M(b)$ has no zero. Let C be a maximal chain in the non-empty set

$$(M(a) \cap M(b)) \cup \bigcap \{L(z) : z \in M(a) \cap M(b)\}$$

and let

$$C^+ = C \cap M(a) \cap M(b),$$
$$C^- = C - C^+.$$

Now C^+ and C^- are non-empty chains, C^+ has no g.l.b., and C^- has no l.u.b. One can show that there exist (generalized) sequences x_α in C^+ and x_β in C^- such that (a) x_α is monotone decreasing and, for each $t \in C^+$, there exists $\alpha(t)$ such that $\alpha > \alpha(t)$ implies $x_\alpha \leqslant t$, and (b) x_β is monotone increasing and, for each $t \in C^-$, there exists $\beta(t)$ such that $\beta > \beta(t)$ implies $x_\beta \geqslant t$. Define $f:X \to C$ as follows: if $x \in \bigcap \{L(x_\alpha)\}$ then

$$f(x) = \min \{x_\beta : x_\beta \not\leqslant x\},$$

and if $x \in X - \bigcap \{L(x_\alpha)\}$ then

$$f(x) = \min \{x_\alpha : x \not\leqslant x_\alpha\}.$$

It is easy to verify that f is well defined and isotone. Further, $f(x_\alpha) < x_\alpha$ and $f(x_\beta) > x_\beta$, so that f is without fixed points. This is a contradiction, whence we infer that $(L(x), \leqslant)$ is a lattice.

LEMMA 3. *Under the hypotheses of Lemma* 2, *if* $x \in X$, *then* $(L(x), \leqslant)$ *is a complete lattice.*

Proof. Let $f:L(x) \to L(x)$ be isotone. Then f can be extended in an isotone manner to $\bar{f}:X \to L(x)$ where

$$\bar{f}(a) = f(a \wedge x).$$

By hypothesis the function $\bar{f}$ has a fixed point which must also be a fixed point of f. By Lemma 2 and Theorem D, $(L(x), \leqslant)$ is a complete lattice.

LEMMA 4. *Under the hypotheses of Lemma* 2, *every maximal chain of* X *is a complete lattice.*

Proof. Let C be a maximal chain of X, and define $f:X \to C$ by

$$f(x) = \sup L(x) \cap C.$$

By Lemma 3, $L(x)$ is a complete lattice for each $x \in X$, and since C meets each $L(x)$, this mapping is well defined, isotone, and $f(x) = x$ if, and only if, $x \in C$. Now if C is incomplete as a lattice then by Theorem D there is an isotone function $g:C \to C$ without fixed points. The composition $gf:X \to C$ is therefore without fixed points, which is a contradiction. Hence $(C, \leqslant)$ is complete as a lattice.

Proof of Theorem 3. The necessity was established in Theorem 2. For the sufficiency, let $(X, \leqslant)$ be a semi-lattice in which every isotone $f:X \to X$ has a fixed point. To prove that X is compact in the interval topology it is sufficient (see the argument of Theorem 1) to prove that if $\mathfrak{F}$ is any non-empty collection of subbasic closed sets with finite intersection property, then $\bigcap \mathfrak{F}$ is non-empty. Now $\mathfrak{F} = \mathfrak{F}_1 \cup \mathfrak{F}_2$ where

$$\mathfrak{F}_1 = \{M(x_\alpha) : \alpha \in A\},$$
$$\mathfrak{F}_2 = \{L(x_\beta) : \beta \in B\}.$$

If $\mathfrak{F}_2$ is non-empty then from Lemma 3 and Corollary 1.1 each $L(x_\beta)$ is compact and hence $\bigcap \mathfrak{F}$ is non-empty. If $\mathfrak{F}_2$ is empty, then $\mathfrak{F}_1$ is not and we may assume that $A = \{\alpha_1, \alpha_2, \dots\}$ is well ordered. Let

$$y_{\alpha_1} = x_{\alpha_1}$$

and, for $\gamma > \alpha_1$,

$$y_\gamma = \inf \bigcap \{M(y_\alpha) : \alpha < \gamma\} \cap M(x_\gamma).$$

To see that y_γ exists, suppose y_α is defined for all $\alpha < \gamma$. Now $\{y_\alpha : \alpha < \gamma\}$ is a chain and hence the set

$$\{z_\alpha : z_\alpha = \inf M(y_\alpha) \cap M(x_\gamma)\}$$

is a chain. By Lemma 4, $z_\gamma = \sup \{z_\alpha : \alpha < \gamma\}$ exists and by Lemma 3, $(L(z_\gamma), \leqslant)$ is a complete lattice so that y_γ exists. Applying Lemma 4 again, $y_0 = \sup \{y_\alpha : \alpha \in A\}$ exists and, clearly, $y_0 \in \bigcap \mathfrak{F}$.

References

1. J. W. Alexander, *Ordered sets, complexes, and the problem of bicompactification*, Proc. Nat. Acad. Sci., *25* (1939), 296–298.
2. Garrett Birkhoff, *Lattice Theory* (rev. ed., New York, 1948).
3. Anne C. Davis, *A characterization of complete lattices*, Pacific J. of Math., *5* (1955), 311–319.
4. O. Frink, *Topology in lattices*, Trans. Amer. Math. Soc. *51* (1942), 569–582.
5. Alfred Tarski, *A lattice-theoretical fixpoint theorem and its applications*, Pacific J. of Math., *5* (1955), 285–309.
6. L. E. Ward, Jr., *Partially ordered topological spaces*, Proc. Amer. Math. Soc. *5* (1954), 144–161.

U.S. Naval Ordnance Test Station
China Lake, California

A CONDITION FOR THE COMMUTATIVITY OF RINGS

I. N. HERSTEIN

A well-known theorem of Jacobson **(1)** asserts that if every element a of a ring A satisfies a relation $a^{n(a)} = a$ where $n(a) > 1$ is an integer, then A is a commutative ring. Thus the condition used in Jacobson's theorem is a sufficient condition for commutativity. However the condition is by no means a necessary one, as it is satisfied by a very restricted class of commutative rings.

In this paper we weaken Jacobson's condition by insisting that it applies only to commutators, and prove that the final result, namely that the ring is commutative, still remains true. In this way, we modify the assumptions used in Jacobson's theorem and produce a condition which is both necessary and sufficient.

The result might be of interest from, possibly, another point of view. The restrictions heretofore used have applied to subrings of the ring whereas the set we consider here is not even an additive subgroup. This suggests a variety of related problems which might be considered. The result may also play a role in the theory of restricted Lie algebras.

We follow the pattern which has become standard by now of ascending from the case of division rings to the general case of arbitrary rings via the Jacobson structure theory.

We begin with

THEOREM 1. *Let D be a division ring in which $(xy - yx)^{n(x,y)} = (xy - yx)$ for all $x,y \in D$ where $n(x,y) > 1$ is an integer. Then D is a commutative field.*

Proof. If $xy - yx = 0$ for all $x,y \in D$ there is, of course, nothing that needs proving. So we assume that for some $a,b \in D$, $ab - ba \neq 0$. Let Z be the center of D. If $\lambda \in Z$, then $\lambda(ab - ba) = (\lambda a)b - b(\lambda a)$, so is again a commutator. Thus by hypothesis

$$(ab - ba)^n = ab - ba, \qquad n > 1, \tag{1}$$

$$[\lambda(ab - ba)]^m = \lambda(ab - ba), \qquad m = m(\lambda) > 1. \tag{2}$$

If we put $S(\lambda) = S = (n - 1)(m - 1) + 1$ then $S > 1$ and we have

$$(ab - ba)^S = (ab - ba) \tag{1.1}$$

$$[\lambda(ab - ba)]^S = \lambda(ab - ba). \tag{2.1}$$

Received November 30, 1956. This paper was supported in part by the ONR contract number SAR/Nonr-609(19) with Yale University.

Since $ab - ba \neq 0$ and since D is a division ring, we deduce from (1.1) and (2.1) that $\lambda^{S(\lambda)} = \lambda$ where $S(\lambda) > 1$ for every $\lambda \in Z$. But then Z must be a field of characteristic $p \neq 0$; moreover, Z is algebraic over its prime field P, which has p elements.

Let $u = ab - ba \neq 0$. Since $u^n = u$, u is algebraic over P, *a fortiori* it is algebraic over Z. Without loss of generality we may assume that $u \notin Z$, for if $u \in Z$ then

$$au = a(ab - ba) = a(ab) - (ab)a$$

is not in Z (for otherwise $a \in Z$ and so $ab - ba = 0$ would follow) and we could carry the argument on for the commutator au rather than for u. Consequently u satisfies a minimal polynomial over Z of degree

$$t > 1, \quad x^t + \lambda_1 x^{t-1} + \ldots + \lambda_t, \qquad \lambda_i \in Z.$$

Let $F = P(\lambda_1, \lambda_2, \ldots, \lambda_t)$ be the field obtained by adjoining $\lambda_1, \lambda_2, \ldots, \lambda_t$ to P. Because the λ_i are algebraic over P and commute with each other, F is a finite field and has, say, q elements. Clearly if $w \in F$ then $w^q = w$. Consider the field $F(u)$. The polynomial $x^q - x$ already has q roots in F, and since it can have at most q roots in $F(u)$, since $u \notin F \subset Z$, we can conclude that $u^q \neq u$. However,

$$u^t + \lambda_1 u^{t-1} + \ldots + \lambda_t = 0$$

so

$$0 = (u^t + \lambda_1 u^{t-1} + \ldots + \lambda_t)^q = u^{qt} + \lambda_1^q u^{q(t-1)} + \ldots + \lambda_t^q$$
$$= (u^q)^t + \lambda_1 (u^q)^{t-1} + \ldots + \lambda_t.$$

Thus u and u^q are both roots of the same minimal polynomial over Z. This implies that there is an element $r \in D$ so that $u^q = rur^{-1}$; that is, $ru = u^q r$. Consequently, $ur \neq ru$ and $(ru - ur)u = u^q(ru - ur)$. Let $y = ur - ru \neq 0$. From the above, $yu = u^q y$. Since y is a commutator, by hypothesis $y^l = y$ for some $l > 1$.

Let

$$T = \left\{ \sum_{i=0}^{l-1} \sum_{j=0}^{n-1} p_{ij} y^i u^j \mid p_{ij} \in P \right\}.$$

T is clearly finite and is an additive subgroup of D; by virtue of $yu = u^q y$, T is also closed under multiplication. Hence T is a finite division ring. By Wedderburn's theorem it follows that T is a commutative field. But both u and y are in T, so $uy = yu$. Since $yu = u^q y$, $uy = yu$ we obtain $u^q = u$, which contradicts $u^q \neq u$. In this way the proof of Theorem 1 is complete.

We recall that a ring A is a prime ring if $aAb = (0)$ implies that either $a = 0$ or $b = 0$. We now proceed to

LEMMA 2. *Let A be a prime ring in which $(xy - yx)^{n(x,y)} = (xy - yx)$, $n(x,y) > 1$. Then A has no non-zero nilpotent elements.*

Proof. If A has nilpotent elements then it has an element $x \neq 0$ such that $x^2 = 0$. If $r \in A$ then $xrx = (xr)x - x(xr)$, so, being a commutator, $(xrx)^n = xrx$ for some $n > 1$. However, $(xrx)^2 \neq xrx^2rx = 0$; whence $0 = (xrx)^n = xrx$. That is, $xAx = (0)$. The primeness of A then forces $x = 0$.

If $e^2 = e$, $e \in A$, it is readily verified that for any $x \in A$, $(xe - exe)^2 = 0$ and $(ex - exe)^2 = 0$. So by Lemma 2 we obtain

LEMMA 3. *If A is as in Lemma 2 then any idempotent in A is in the center of A.*

We now go to the next step in the Jacobson structure theory approach and prove

THEOREM 4. *If A is a primitive ring in which $(xy - yx)^{n(x,y)} = (xy - yx)$ for all $x,y \in A$ where $n(x,y) > 1$ is an integer, then A is a commutative field.*

Proof. Since A is a primitive ring it possesses a maximal right ideal ρ which contains no non-zero two-sided ideal of A. Thus $\rho \cap Z = (0)$ (where Z is the center of A) for if $x \in \rho \cap Z$ then $xA = Ax \subset \rho$ is a two-sided ideal of A which is located in ρ, so must be (0); by the primitivity of A we must conclude that $x = 0$.

Let $x,y \in \rho$. By the hypothesis, for some $n > 1$, $(xy - yx)^n = (xy - yx)$. But then $e = (xy - yx)^{n-1} \in \rho$ is an idempotent, so it must be in Z by Lemma 3. That is $e \in \rho \cap Z$. By the above remarks this implies that $e = 0$; thus

$$0 = e(xy - yx) = (xy - yx)^n = xy - yx.$$

That is, any two elements of ρ commute with each other. Suppose $a, b \in \rho$ and $r \in A$. Since $ar \in \rho$, $(ar)b = b(ar)$. However, $ab = ba$, so we deduce that $a(br - rb) = 0$ for all $a,b \in \rho$, $r \in A$. Thus $\rho(br - rb) = (0)$, which, in a primitive ring, means that either $\rho = 0$ or $br - rb = (0)$. Thus $b \in Z$, whence $b \in \rho \cap Z$ from which, as before, $b = 0$. But then $\rho = (0)$ is a maximal right ideal in the primitive ring A; in consequence A must be a division ring, which, by Theorem 1, must in turn be a commutative field.

If A is a ring semi-simple in the sense of Jacobson then A is isomorphic to a subdirect sum of primitive rings. Each of these primitive rings is a homomorphic image of A, and so inherits the property that

$$(xy - yx)^{n(x,y)} = (xy - yx).$$

By Theorem 4 these primitive rings must all be commutative fields, and so we have

THEOREM 5. *If A is a semi-simple ring in which $(xy - yx)^{n(x,y)} = (xy - yx)$ for all $x,y \in A$ then A is commutative.*

We now have all the preliminaries needed to prove the main theorem of this paper, namely

THEOREM 6. *Let A be a ring in which $(xy - yx)^{n(x,y)} = (xy - yx)$ for all $x,y \in A$ where $n(x,y) > 1$ is an integer. Then A is a commutative ring.*

Proof. Let N be the radical of A. Hence A/N is semi-simple, and so, by Theorem 5, it is commutative. Thus $xy - yx \in N$ for all $x,y \in A$. However, $(xy - yx)^n = (xy - yx)$, so $e = (xy - yx)^{n-1}$ is an idempotent; moreover $e \in N$. But the only idempotent in the radical is 0. So $(xy - yx)^{n-1} = 0$ from which $0 = (xy - yx)^n = (xy - yx)$. Thus A is commutative.

REFERENCES

1. N. Jacobson, *Structure theory for algebraic algebras of bounded degree*, Ann. Math., *46* (1945), 695–707.

Yale University

ON THE STRUCTURE OF FROBENIUS GROUPS

WALTER FEIT

1. Introduction. Let G be a group which has a faithful representation as a transitive permutation group on m letters in which no permutation other than the identity leaves two letters unaltered, and there is at least one permutation leaving exactly one letter fixed. It is easily seen that if G has order mh, a necessary and sufficient condition for G to have such a representation is that G contains a subgroup H of order h which is its own normalizer in G and is disjoint[1] from all its conjugates. Such a group G is called a Frobenius group of type (h, m).

Some immediate consequences of the definition are that in a Frobenius group G of type (h,m), h divides $m-1$, every element other than the identity whose order divides h is contained in exactly one subgroup of order h, and any two subgroups of order h are conjugate. A fundamental property of a Frobenius group G of type (h,m) is that G contains exactly m elements whose order divides m and these form **(2**, p. 334**)** a normal subgroup M of G. This subgroup M will be called the regular subgroup of G, since the above mentioned permutation representation of G when restricted to M is just the regular representation of M.

Burnside has shown that the regular subgroup of a Frobenius group of type (h, m), with even h, is abelian of odd order **(2**, p. 172**)**. Conversely it is not hard to show that an abelian group of odd order m can be imbedded in a Frobenius group G of type $(2, m)$ as the regular subgroup of G. In general, the regular subgroup of a Frobenius group need not be abelian (see **4** for a counter example), however it has been conjectured that it must always be nilpotent.[2] The main result proved below is that, under certain conditions, the regular subgroup of a Frobenius group is nilpotent. If it can be shown that no exceptional groups exist (in the sense of §4), then the nilpotency would be proved in general. The result can also be restated in a different form using the language of automorphisms[3] as is done in the Corollary in §4.

2. Some properties of Frobenius groups

LEMMA 2.1. *Let G be a group of order hm, where h and m are relatively prime and unequal to* 1. *Then G is a Frobenius group of type (h, m) if and only if G*

Received April 8, 1957.

[1]Two subgroups of a group are said to be disjoint if their intersection consists of only the identity element.

[2]I am indebted to Professor Marshall Hall for telling me about this conjecture.

[3]After I had written up this paper I was informed by Professor Herstein that he and Professor Wielandt had proved a result essentially equivalent to the Corollary in §4, though their methods were somewhat different from those used here.

contains a normal subgroup M of order m, and the order of every element divides either h or m.

Proof. If G is a Frobenius group of type (h, m), then its regular subgroup M is normal in G and has order m. Since h and m are relatively prime, every element x in G can be written as a product $x = x_1x_2$, where x_1 and x_2 commute and $x_1^h = 1 = x_2^m$. If $x_1 \neq 1$, then it lies in some subgroup H of order h, hence

$$x_2^{-1}x_1x_2 = x_1 \in H \cap x_2^{-1}Hx_2.$$

Since distinct subgroups of order h are disjoint, x_2 must lie in the normalizer of H, and as H is its own normalizer, x_2 must lie in H; therefore $x_2 = 1$. Consequently either $x_1 = 1$ or $x_2 = 1$, and the order of every element in G must divide either h or m.

Conversely, as the index of M in G is relatively prime to m, G contains a subgroup H of order h (**5**, p. 132, Theorem 25) and every element whose order divides h lies in a subgroup conjugate to H (**3**, p. 184, Lemma 6.1). Let N be the normalizer of H, $M \cap N$ is a normal subgroup of N, hence N is the direct product of H and $M \cap N$, thus if $M \cap N \neq \{1\}$, N contains elements where order divides neither h nor m which is impossible, therefore $N = H$, and G contains m subgroups of order h conjugate to H. Each of these subgroups contains $h - 1$ elements other than the identity, if some element $x \neq 1$ is contained in two of these subgroups, then the total number of elements unequal to the identity whose order divides h is strictly less than $m(h - 1) = mh - m$. All the elements whose order divides m lie in M and thus there are exactly m of them, hence the number of elements in G, other than the identity, whose order divides h must equal $mh - m$. Therefore the assumption that G contains two subgroups of order h which are not disjoint is untenable and the proof is complete.

LEMMA 2.2. *Let G be a Frobenius group of type (h, m). If G_1 is a subgroup of order h_1m_1, where h_1 divides h, m_1 divides m, and $h_1 \neq 1 \neq m_1$, then G_1 is a Frobenius group of type (h_1, m_1).*

Proof. By Lemma 2.1, every element of G_1 has order dividing h_1 or m_1. If M is the regular subgroup of G, then $M \cap G_1$ is a normal subgroup of G_1 whose order is m_1, hence Lemma 2.1 implies the result.

LEMMA 2.3. *Let G be a Frobenius group of type (h, m). If $\bar{G}$ is a homomorphic image of G whose order is hm_1, with $m_1 > 1$, then $\bar{G}$ is a Frobenius group of type (h, m_1).*

Proof. Let K of order k be the kernel of the homomorphism, then $K \in M$, and the image $\bar{M}$ of M is a normal subgroup of $\bar{G}$ whose order is $m/k = m_1$.

If $\bar{x}$ is an element of $\bar{G}$ and x is some element of G which is mapped onto $\bar{x}$, then $x^n = 1$ implies that $\bar{x}^n = 1$. Since the order of x divides either h or m, this is also the case for $\bar{x}$. If the order of $\bar{x}$ divides m, then it must divide the

greatest common divisor m_1 of m and the order m_1h of $\bar{G}$. Hence the order of every element in $\bar{G}$ divides either h or m_1 and Lemma 2.1 now implies that $\bar{G}$ is a Frobenius group of type (h_1, m_1).

LEMMA 2.4. *Suppose G is a Frobenius group of type (h, m), let r be a prime dividing h, then G contains a subgroup G_1 which is a Frobenius group of type (r, m). The regular subgroup of G is also the regular subgroup of G_1.*

Proof. Let M be the regular subgroup of G. G contains a subgroup R of order r. Since M is normal in G, $G_1 = MR$ is a group of order mr which contains M as a normal subgroup and in which the order of every element divides either m or r, hence Lemma 2.1 yields the desired result.

LEMMA 2.5. *Suppose that G is a Frobenius group of type (h, m). Let p be a prime dividing m and let T be a subgroup of some Sylow p-group P of G which is normal in the normalizer $N(P)$ of P. If the normalizer $N(T)$ of T has order n, then h divides n and $N(T)$ is a Frobenius group of type $(h, n/h)$.*

Proof. As the regular subgroup M of G is normal in G and its order m is divisible by the highest power of p which divides the order of G, all Sylow p-groups lie in M and hence are conjugate in M. Therefore the number of Sylow p-groups (in G as well as in M) is a divisor of m, then the index of the normalizer $N(P)$ of P in G divides m, hence h divides the order of $N(P)$. By assumption T is normal in $N(P)$, therefore h divides the order n of $N(T)$. By Lemma 2.2 $N(T)$ is a Frobenius group of the desired type.

LEMMA 2.6. *Let G be a Frobenius group of type (h,m). Suppose that A is a subgroup of G of order q^2 or qr, where q and r are primes dividing h, then A is cyclic.*

Proof. Let M be the regular subgroup of G. As M is normal in G, MA is a group and hence by Lemma 2.2 a Frobenius group whose regular subgroup is M. Let p be a prime dividing m and let P be a Sylow p-group of M. Let C be the subgroup of P consisting of all elements in the center of P whose pth power is the identity. Clearly C is a characteristic subgroup of P hence normal in $N(P)$, therefore by Lemma 2.5 $N(C)$ is a Frobenius group. The group $N(C)$ contains a Frobenius group $G_1 = A_1C$ whose regular subgroup is C, and which contains a subgroup A_1 conjugate to A.

Each element x in A_1 defines an automorphism of C which sends y into $x^{-1}yx$ for y in C, hence A_1 can be considered to be a group of automorphisms of C. As A_1C is a Frobenius group with regular subgroup C, no element of A_1 commutes with any element of C. In other words, no automorphism of A_1 leaves any element of C other than the identity fixed. An argument of Burnside[4] can now be applied which shows that A_1 is cyclic, as A is conjugate to A_1, it too must be cyclic.

[4]Burnside deduced a false theorem from a correct argument; this can be found in (2, pp. 334–335). For a statement and proof of the result, in the form needed above, see (6, p. 196).

3. The structure of a special class of groups. This section is devoted to investigating groups satisfying certain assumptions. In order to prevent repetition, the basic hypothesis will be stated separately. The symbols $\otimes$ and $\oplus$ will stand for direct product and direct sum respectively. For any subset S of G, the centralizer of S in G will be denoted by $C(S)$.

Hypothesis I. *The order of G is $p^a q^b$, where p and q are distinct primes and $a,b > 0$. The Sylow p-group P of G is the direct product of groups of order p and is normal in G. A Sylow q-group of G is the direct product of groups of order q.*

LEMMA 3.1. *Suppose that G satisfies hypothesis* I. *There is a one to one mapping from P onto an a-dimensional vector space V over the field F of p elements with the property that $\overline{y_1 y_2} = \bar{y}_1 + \bar{y}_2$ for y_1, y_2 in P, where $\bar{y}$ denotes the image of y in V. Let Q be a Sylow q group of G, for x in Q define the linear transformation $A(x)$ acting on V by $A(x)\bar{y} = \overline{xyx^{-1}}$ for all y in P. The mapping of Q into the group of linear transformations on V defined in this way is a completely reducible representation Γ of Q. A subgroup of P is normal in G if and only if the corresponding subspace $\bar{P}$ of V is invariant under the representation Γ.*

Proof. Every statement of the Lemma can easily be checked. The complete reducibility of Γ follows from the fact that the characteristic p of F does not divide the order q^b of Q.

LEMMA 3.2. *If G satisfies hypothesis* I, *and if P contains a subgroup T which is normal in G, then G contains a normal subgroup T_1 with the property that $P = T \otimes T_1$.*

Proof. Let $\bar{T}$ be the subspace of V corresponding to T under the mapping defined in Lemma 3.1. Since T is normal in G, $\bar{T}$ is invariant under the representation Γ. The complete reducibility of Γ implies the existence of an invariant subspace $\bar{T}_1$ such that $V = \bar{T} \oplus \bar{T}_1$. Hence by Lemma 3.1, T_1 is a normal subgroup of G and $P = T \otimes T_1$.

LEMMA 3.3. *Suppose that G satisfies hypothesis* I, *let Q be a Sylow q-group of G. If x is in Q, then $C(x) \cap P$ is a normal subgroup of G. If P_0 is a minimal normal subgroup of G which is contained in P, then q^{b-1} divides the order of $C(P_0)$.*

Proof. Since P is a normal subgroup of G, $C(x) \cap P$ is a normal subgroup of $C(x)$. Therefore the normalizer of $C(x) \cap P$ contains $C(x)$ which contains Q, since x lies in the abelian group Q. As P is abelian, it lies in the normalizer of every subgroup. Hence $G = PQ$ is contained in the normalizer of $C(x) \cap P$.

The group $C(x) \cap P_0$ is a normal subgroup of G since it is the intersection of the normal subgroups $C(x) \cap P$ and P_0. Therefore since P_0 is a minimal normal subgroup of G, either P_0 is contained in $C(x)$ or disjoint from it. In other words, every element x of Q either lies in the centralizer of P_0 or commutes with no element of P_0 other than the identity.

Suppose that q^{b-1} does not divide the order of $C(P_0)$, then Q contains a subgroup Q_0 of order q^2 which is disjoint from $C(P_0)$. As P_0 is normal in G, $G_0 = P_0Q_0$ is a group containing P_0 as a normal subgroup. Since Q_0 is a Sylow q-group of G_0, every element whose order divides q^2 is conjugate to an element of Q_0, and hence commutes with no element of order p. Therefore the order of every element divides either q^2 or p^{a_0}, where p^{a_0} is the order of P_0. Lemma 2.1 now implies that G_0 is a Frobenius group of type

$$(q^2, p^{a_0}),$$

consequently Q_0 is cyclic by Lemma 2.6. This is impossible since Q_0 contains no element of order q^2. Thus the assumption that q^{b-1} does not divide the order of $C(P_0)$ has led to a contradiction, which proves the result.

LEMMA 3.4. *Let G be a group which satisfies hypothesis* I. *Assume that there exists an automorphism σ of G of prime order r which sends some Sylow q group Q of G into itself. Furthermore suppose that $\sigma(s) \neq s$ where s is any proper subgroup of Q, or any proper subgroup of P which is normal in G, or any element of G other than the identity. Then either G is abelian*[5] *or $b = 1$.*

Proof. Suppose that G is not abelian and $b > 1$. Let P_1 be a minimal normal subgroup of G, if q^b divides the order of $C(P_1)$, then P_1 lies in the center of G, hence the p-Sylow group of the center of G must equal P since it is mapped into itself by σ, but then G is abelian contrary to assumption. Hence the Sylow q-group of $C(P_1)$ has order q^{b-1} by Lemma 3.3. By taking a group conjugate to P_1 if necessary, it may be assumed that there is a q-Sylow group Q_1 of $C(P_1)$ which is contained in Q.

Define $P_{i+1} = \sigma(P_i)$ for all i, let k be the smallest integer with the property that

$$\bar{P}_{k+1} \subset \bar{P}_1 + \ldots + \bar{P}_k,$$

where the bar denotes the mapping defined in Lemma 3.1. If y_1 is any element in P_1, and $y_{i+1} = \sigma(y_i)$, then σ maps the product $y_1 \ldots y_r$ into itself, hence by assumption this product is 1, and therefore y_r is contained in $P_1 \ldots P_{r-1}$. As y_1 was chosen arbitrarily in P_1, this states that $P_r \subset P_1 \ldots P_{r-1}$, and hence $\bar{P}_r \subset \bar{P}_1 + \ldots + \bar{P}_{r-1}$, consequently $k < r$. Since σ maps $P_1 \ldots P_k$ into itself and no proper subgroup of P which is normal in G has this property, $P = P_1 \ldots P_k$, and therefore $\bar{P} = \bar{P}_1 + \ldots + \bar{P}_k$. The representation is completely reducible, hence (**1**, Theorem 1.4C) $\bar{P} = \bar{P}_1 \oplus \ldots \oplus \bar{P}_k$.

Let $Q_i = C(P_i) \cap Q$, it is easily seen that $Q_{i+1} = \sigma^i(Q_1)$, hence by the choice of P_1, the order of Q_i is exactly q^{b-1}. Since $b > 1$, each group Q_i is a proper subgroup of Q, therefore $\sigma(Q_i) \neq Q_i$ for all i. As $k < r$, this implies that $Q_i \neq Q_j$ for $1 \leqslant i < j \leqslant k$, because otherwise $\sigma^{j-i}(Q_i) = Q_i$ which in turn implies $\sigma(Q_i) = Q_i$, as $0 < j - i < k < r$ and $j - i$ and r are relative prime.

[5] Actually G is abelian in all cases, this is a consequence of the Theorem below.

Suppose that there exists an operator isomorphism ρ from the Γ module $\bar{P}_i$ onto the Γ module $\bar{P}_j$ with $1 \leqslant i < j \leqslant k$. As Q_i and Q_j are distinct subgroups of Q of order q^{b-1} it is possible to find an element x in Q_j which is not contained in Q_i. Then

$$\rho\{A(x)\bar{y}\} = A(x)\{\rho(\bar{y})\}$$

for y in $\bar{P}_i$. Since x is in Q_j, $\rho(y)$ is in $\bar{P}_j$, $A(x)\{\rho(\bar{y})\} = \rho(\bar{y})$, therefore $\rho\{A(x)\bar{y}\} = \rho(\bar{y})$ which implies that $A(x)\bar{y} = \bar{y}$ which finally yields that x is in Q_i contrary to the choice of x. Consequently the Γ module $\bar{P}_i$ is not operator isomorphic to the Γ module $\bar{P}_j$ for $1 \leqslant i < j \leqslant k$. This implies that the irreducible subspaces of $\bar{P}$ are unique (**1**, Theorem 1.6C), in other words the only irreducible subspaces of $\bar{P}$ are $\bar{P}_1, \ldots, \bar{P}_k$, hence the only minimal normal subgroups of G which are contained in P are $P_1, \ldots, P_k$. As σ is an isomorphism of G mapping P into itself, $\sigma(P_k)$ is a minimal normal subgroup of G contained in P, therefore $\sigma(P_k) = P_i$ for some i between 1 and k, hence $\sigma^{k+1-i}(P_i) = P_i$. As $0 < k + 1 - i \leqslant k < r$, $k + 1 - i$ and r must be relatively prime and $\sigma(P_i) = P_i$, therefore $\sigma(Q_i) = Q_i$, which was shown to be impossible. Hence the assumption that G is non-abelian and $b > 1$ has led to a contradiction which proves the Lemma.

4. The regular subgroup of a Frobenius group. Before proceeding to investigate the structure of the regular subgroup of a Frobenius group it is necessary to make the following definition.

Definition. A group G is said to be *exceptional* if G is a non-cyclic simple group in which the normalizer of every characteristic subgroup $\neq \{1\}$ of a Sylow p-group P of G is P, for all primes p dividing the order of G.

No known groups are exceptional in the sense defined above. A special case of a conjecture of Zassenhaus (**7**, footnote on p. 6) would be sufficient to prove that exceptional groups do not exist. The case treated in the theorem below is concerned with groups in which no subgroup has a composition factor which is an exceptional group. This is a large class of groups as is shown by the following Lemma.

LEMMA 4.1. *If G is a solvable group, or if every Sylow group of G is abelian, then no subgroup of G has a composition factor which is an exceptional group.*

Proof. If G is solvable, the result is immediate as no simple group can occur as a composition factor of a subgroup. If every Sylow group of G is abelian then this is also the case for every composition factor of a subgroup, hence it suffices to show that a group H in which every Sylow group is abelian cannot be exceptional. Let P be a Sylow p-group of H for some prime p dividing the order of H. Suppose that P is its own normalizer, then by a theorem of Burnside (**5**, p. 139), H cannot be simple and consequently not exceptional.

LEMMA 4.2. *Let M be the regular subgroup of a Frobenius group G of type (h,m). Suppose that M is not a direct product of exceptional groups, and that the regular subgroup of every proper subgroup of G which is a Frobenius group of type (h,m_1) is nilpotent, then M contains a normal subgroup of prime power order.*

Proof. Let H be a subgroup of G of order h. If M contains a proper characteristic subgroup M_1 of order m_1, then M_1 is normal in G, hence by Lemma 2.2, M_1H is a Frobenius group of type (h,m_1), consequently M_1 is nilpotent and any Sylow subgroup of M_1 is normal in M.

Suppose now that M is characteristically simple and contains no normal subgroup of prime power order. Then M is the direct product of isomorphic non-cyclic simple groups $M_1, \ldots, M_s$. By assumption these are not exceptional therefore there is a prime p such that the Sylow p-group P_1 of M_1 contains a characteristic subgroup T_1 such that $N(T_1) \cap M_1 \neq P_1$, where $N(T_1)$ denotes the normalizer of T_1 in G. Let P_i and T_i denote the images in M_i of P_1 and T_1 respectively, under an isomorphism mapping M_1 onto M_i. Let

$$P = P_1 \otimes \ldots \otimes P_s, \qquad T = T_1 \otimes \ldots \otimes T_s;$$

it is clear then that P is a Sylow p-group of M, T is a normal subgroup of $N(P)$ and $P \neq N(T) \cap M$. By assumption T is not normal in M, hence $N(T) \neq G$, then by Lemma 2.5 and the assumption of this Lemma $N(T) \cap M$ is nilpotent, hence P is a normal subgroup of $N(T) \cap M$ consequently $P \neq N(P) \cap M$. Let C be the center of P, C is a normal subgroup of $N(P)$, therefore $P \neq N(C) \cap M$. As $N(C) \neq G$ Lemma 2.5 once again yields that $N(C) \cap M$ is nilpotent, hence the p-commutator subgroup of $N(C) \cap M$ is a proper subgroup of $N(C) \cap M$. If it can be established that $N(C) \cap M$ is p-normal, a result of Grün **(5**, Theorem 6, p. 141**)** will imply that $M \neq M'$, where M' is the commutator subgroup of M. As M is characteristically simple this yields that $M' = \{1\}$, hence M is abelian in contradiction with the assumption that M contains no normal subgroup of prime power order, and the Lemma is proved. We now proceed to show that M is p-normal.

Suppose $C \subset xPx^{-1}$ for some x, then $x^{-1}Cx \subset P$. As $N(C) \cap M \neq P$, there is a prime q different from p which divides the order of $N(C) \cap M$, let Q be a Sylow q-group of $N(C) \cap M$. As $N(C) \cap M$ is nilpotent and contains both P and Q they commute elementwise. Since $x^{-1}Cx \subset P$, Q commutes elementwise with $x^{-1}Cx$ and therefore is contained in $N(x^{-1}Cx)$. Since $x^{-1}Px$ and Q are contained in the nilpotent group $N(x^{-1}Cx) \cap M$ they also commute elementwise, hence P and $x^{-1}Px$ are both contained in $N(Q)$. As Q is characteristic in $N(C) \cap M$ it is normal in $N(C)$, hence h divides the order of $N(Q)$, as $N(Q) \neq G$, the assumptions of the Lemma yield that $N(Q) \cap M$ is nilpotent, consequently $P = x^{-1}Px$ as both P and $x^{-1}Px$ are Sylow p-groups of a nilpotent group. Therefore the center C of P is contained in no other Sylow p-groups of M, hence M is p-normal, which suffices to prove the Lemma.

THEOREM. *Let M be the regular subgroup of a Frobenius group G. Suppose that no subgroup of M has an exceptional group as a composition factor, then M is nilpotent.*

Proof. Suppose that the theorem is false. Let M of order m be a non-nilpotent group of minimum order in which no subgroup has an exceptional group as a composition factor, and which can be represented as the regular subgroup of some Frobenius group G. Pick a prime r which divides the order of G but does not divide m and let R be a subgroup of G of order r, then $G_0 = RM$ is a Frobenius group of type (r,m) by Lemma 2.2. Suppose M has a non-trivial center C, then C is characteristic in M and therefore normal in G_0. It is clear that M/C satisfies the assumption of the theorem and has order less than m, hence M/C is nilpotent. It follows directly from the definition of a nilpotent group that this implies that M is nilpotent contradicting the choice of M. Thus the center of M is $\{1\}$.

As a first step in the proof it will be shown that $m = p^a q^b$, where p and q are primes and where the Sylow p-group of M is normal in M. The group M satisfies the assumption of Lemma 4.2, hence for some prime p dividing m, M contains a normal subgroup whose order is a power of p. Let P be a maximal normal subgroup of M whose order is a power of p, then P is characteristic in M and hence normal in G_0. By Lemma 2.3, G/P is a Frobenius group whose regular subgroup is M/P. It is clear that M/P satisfies the assumptions of the theorem and has order less than m, hence M/P is nilpotent. The Sylow p-group of M/P is a normal subgroup of M/P, hence its inverse image in M is normal in M and contains P, it follows from the way P was chosen that P is the Sylow p-group of M. Let $q_1, \ldots, q_s$ be the distinct primes, other than p, which divide m, let Q_i be a Sylow q_i-group of M. Since M/P is nilpotent, Q_iP/P is normal in M/P, hence Q_iP is normal in M, therefore Q_iP is characteristic in M and hence normal in G_0. Consequently, Lemma 2.2. implies that Q_iPR is a Frobenius group whose regular subgroup is Q_iP. If $s > 1$, $Q_iP \neq M$, hence Q_iP is nilpotent, therefore every element of P commutes with every element of Q_i. Since this is the case for all i, the center of P lies in the center of M, which leads to a contradiction since M has no non-trivial center. Therefore[6] $s = 1$ and $m = p^a q^b$.

As all the Sylow q-groups of G lie in M and are conjugate in M, the index of the normalizer $N(Q)$ of a Sylow q-group Q of G divides m, therefore r divides the order of $N(Q)$. Let R_0 be a subgroup of $N(Q)$ of order r, the group Q_0 consisting of all elements in the center of Q whose order divides q is a characteristic subgroup of Q, hence $R_0 \subset N(Q_0)$, therefore R_0Q_0 is a group and also R_0Q_0P is a group. By Lemma 2.2 this is a Frobenius group whose regular subgroup is Q_0P. If $Q_0 \neq Q$, then $Q_0P \neq M$, hence Q_0P is nilpotent, therefore both P and Q lie in $C(Q_0)$, thus Q_0 is in the center of M which is impossible,

[6] $s < 1$, hence either $s = 1$ or $s = 0$, in the latter case M is a p-group, which is impossible, hence only the case $s = 1$ needs to be considered.

hence $Q = Q_0$. Let P_0 be the group consisting of all the elements in the center of P whose order divides p, then as before R_0QP_0 is a Frobenius group whose regular subgroup is QP_0, if $P_0 \neq P$, then $QP_0 \neq M$ and QP_0 is nilpotent, hence P_0 lies in the center of M which is impossible. Therefore $P = P_0$. Consequently the group G satisfies hypothesis I of section 3.

Pick an element x in R, then the mapping $\sigma(y) = xyx^{-1}$ defines an automorphism σ of M of prime order r with the property that $\sigma(y) \neq y$ for all $y \neq 1$ in M. We wish to show that M satisfies the hypothesis of Lemma 3.4. Suppose $Q_0 \neq \{1\}$ is a subgroup of Q such that $\sigma(Q_0) = Q_0$, then $R \subset N(Q_0)$, therefore RQ_0 is a group and by Lemma 2.2 RQ_0P is a Frobenius group whose regular subgroup is Q_0P. If $Q \neq Q_0$, then $Q_0P \neq M$, hence Q_0P is nilpotent, therefore Q_0 lies in the center of M (as Q is abelian) which is impossible, therefore $Q = Q_0$. Suppose $P_0 \neq \{1\}$ is subgroup of P which is normal in M such that $\sigma(P_0) = P_0$, then $R \subset N(P_0)$, therefore P_0 is normal in G, hence RQP_0 is a group and Lemma 2.2 can once again be applied to show that QP_0 is nilpotent if $P_0 \neq P$, this leads to the fact P_0 is contained in the center of M which cannot be the case and P_0 must equal P. Therefore M satisfies the assumption of Lemma 3.4, hence, that Lemma implies that the order of Q is q, since M was assumed to be non-abelian.

If any element x in P commutes with any element of order q, then $C(x)$ contains P and is divisible by q, therefore x lies in the center of M, hence $x = 1$. In other words, the order of every element of M divides either p^e or q, hence the order of every element in G divides either p^e or q or r, since no element of order r commutes with any element whose order is not r. Therefore every element of G has an order dividing p^e or qr and by Lemma 2.1, G is a Frobenius group of type (qr,p^e), consequently Lemma 2.6 implies that QR is a cyclic group. This is impossible since G contains no elements of order qr. The assumption that the Theorem is false has led to a contradiction and the proof is complete.

COROLLARY. *Let M be a group which admits a group of automorphisms A in which no automorphism other than the identity leaves any element of M other than the identity invariant. Furthermore assume that no subgroup of M has an exceptional group as a composition factor, then M is nilpotent.*

Proof. Let G be the group defined by extending M by A (**5**, pp. 94-98), then it is easily seen that G is a Frobenius group whose regular subgroup is M, hence Theorem 1 yields the desired result.

References

1. E. Artin, C. J. Nesbitt and R. M. Thrall, *Rings with Minimum Condition* (Ann Arbor, 1948).
2. W. Burnside, *Theory of Groups of Finite Order* (Dover, New York, 1955).
3. W. Feit, *On a conjecture of Frobenius*, Proc. Amer. Math. Soc., *7* (1956), 177–188.
4. O. J. Schmidt, *Ueber die Frobenius Gruppen*, C.R. (Doklady) Acad. Sci. URSS (N.S.), *26* (1940), 3–5.
5. H. Zassenhaus, *The theory of Groups* (Chelsea, New York, 1949).
6. ———, *Ueber endliche Fastkörper*, Abh. Math. Sem. Hamburg Univ., *11* (1935), 187–220.
7. ———, *Ueber Liesche Ringe mit Primzahlcharacteristik*, Abh. Math. Sem. Hamburg Univ., *13* (1940), 1–100.

Cornell University

FACTORIZATION RINGS

J.-M. MARANDA

1. Introduction. Let $\mathfrak{o}$ be an integral domain with $\mathfrak{K}$ as field of quotients. W. Krull has shown **(3; 4)** that the following three conditions on $\mathfrak{o}$ are equivalent:

1. There is a set of rank 1, discrete valuations of $\mathfrak{K}$, $\{V_i\}_{i \in I}$, such that for each non-null element $a \in \mathfrak{K}$, $V_i(a) = 0$ for all $i \in I$ except a finite number, and such that for all $a \in \mathfrak{K}$, $a \in \mathfrak{o}$ if and only if $V_i(a) \geqslant 0$ for all $i \in I$.

2. Every non-trivial principal ideal of $\mathfrak{o}$ is the intersection of a finite number of formal powers of minimal non-trivial prime ideals of $\mathfrak{o}$.

3. The partially ordered semi-group of classes of quasi-equal non-null ideals (fractional) of $\mathfrak{o}$ is a group with unique factorization theorem.

Krull called an integral domain $\mathfrak{o}$ satisfying these conditions an "endliche diskrete Hauptordnung" and showed that there is a minimum set of rank 1, discrete valuations of $\mathfrak{K}$ satisfying 1. We may notice that a Dedekind ring is an "endliche diskrete Hauptordnung" for which the theory of quasi-equality is trivial, i.e. if two ideals are quasi-equal, then they are equal.

The object of this paper is to generalize this theory of integral domains to a theory of arbitrary commutative rings with unity element. For simplicity, we will call these generalized "endliche diskrete Hauptordnungen" "factorization rings".

The reader will soon realize that the theory of quasi-equality of van der Waerden and Artin **(7, §105)**, generalized to the case of an arbitrary commutative ring with unity element, is the fundamental tool utilized.

We will obtain in particular, those known results concerning Noetherian rings that are integrally closed in their full ring of quotients, that are given in **(5, §4.7 and §4.9)**, most of them in a more general context (the ascending chain condition is not necessarily valid for the ideals of a factorization ring), and by methods that are undoubtedly "multiplicative."

In §6 we will define the notion of a "generalized Dedekind ring", and although we cannot go into any details here in the introduction, we may remark that for such a generalized Dedekind ring $\mathfrak{o}$, if the ascending chain condition is valid for its ideals, then, for any ideal $\mathfrak{a}$ of $\mathfrak{o}$,

$$\mathfrak{a} = \mathfrak{a}_S \mathfrak{p}_1^{m_1} \mathfrak{p}_2^{m_2} \dots \mathfrak{p}_r^{m_r}$$

where S is the set of all regular elements of $\mathfrak{o}$, where $\mathfrak{a}_S$ is the isolated component of $\mathfrak{a}$ determined by S and where the $\mathfrak{p}_i$ are relevant prime ideals **(5, p. 76)** of $\mathfrak{o}$, and this decomposition is "unique" in a certain sense. This is an

Received August 15, 1956; in revised form October 2, 1956.

obvious generalization of the unique decomposition theorem for the ideals of an ordinary Dedekind ring.

Finally, if $\mathfrak{O}$ is a commutative ring with unity element, and if the descending chain condition is valid for the ideals of $\mathfrak{O}$, then we will determine, in §7, all orders of $\mathfrak{O}$ that are Noetherian generalized Dedekind rings.

2. Valuations and Subvaluations. Let us consider a commutative and associative ring $\mathfrak{O}$.

Definition. A function V of $\mathfrak{O}$ onto a partially ordered semi-group M will be called a *valuation*[1] of $\mathfrak{O}$ if for all $a, b, c \in \mathfrak{O}$,

V1. $V(a) \leqslant V(b) \,\&\, V(a) \leqslant V(c) \rightarrow V(a) \leqslant V(b - c)$

V2. $V(ab) = V(a)V(b)$

We will call M the ordered semi-group of values of V. In the case where M is totally ordered, V1 may be replaced by

V1′. $V(b - c) \geqslant \min\{V(b), V(c)\}$

Definition. A reflexive and transitive binary relation R on $\mathfrak{O}$ will be called a *subvaluation* of $\mathfrak{O}$ if for all $a, b, c, d \in \mathfrak{O}$,

S1. $aRb \,\&\, aRc \rightarrow aR(b - c)$

S2. $aRb \,\&\, cRd \rightarrow acRbd$

If V is a valuation of $\mathfrak{O}$ and if we define the relation R on $\mathfrak{O}$ as follows: for all $a, b \in \mathfrak{O}$, aRb if and only if $V(a) \leqslant V(b)$, then it is easily verified that R is a subvaluation of $\mathfrak{O}$. We will say that R is the subvaluation of $\mathfrak{O}$ determined by V.

Conversely, let R be a subvaluation of $\mathfrak{O}$. If we define the relation $\bar{R}$ on $\mathfrak{O}$ as follows: for all $a, b \in \mathfrak{O}$, $a\bar{R}b$ if and only if aRb and bRa, then one can verify the following:

1. The relation $\bar{R}$ is an equivalence relation and if V is the natural function of $\mathfrak{O}$ onto the quotient set $M = \mathfrak{O}/\bar{R}$, then one may define a partial ordering relation on M as follows: for all $a, b \in \mathfrak{O}$, $V(a) \leqslant V(b)$ if and only if aRb.

2. The relation $\bar{R}$ is multiplicative so that one can define an induced multiplication on M, and with respect to this operation and the partial ordering relation defined above, M is a partially ordered semi-group.

3. The function V is a valuation of $\mathfrak{O}$ with M as ordered semi-group of values and V determines the given subvaluation R of $\mathfrak{O}$.

Let V be a valuation of $\mathfrak{O}$ with ordered semi-group of values M and let R be the subvaluation of $\mathfrak{O}$ determined by V. For all $a \in \mathfrak{O}$, we have $aR0$ and $aR(-a)$, for, since R is reflexive, aRa, and by S1,

$$aRa \,\&\, aRa \rightarrow aR(a - a)$$
$$aR0 \,\&\, aRa \rightarrow aR(0 - a)$$

[1] In the case where $\mathfrak{O}$ is a field and M is a partially ordered group with added symbol ∞, this definition is not new, see e.g. **(2)**.

The corresponding properties for V are: for all $a \in \mathfrak{O}$, $V(a) \leqslant V(0)$ and $V(a) \leqslant V(-a)$ so that $V(a) = V(-a)$.

PROPOSITION 1. *If*

$$\mathfrak{a} = \{a \in \mathfrak{O} | 0Ra\} = \{a \in \mathfrak{O} | V(a) = V(0)\}$$

then $\mathfrak{a}$ *is an ideal of* $\mathfrak{O}$ *which we will call the kernel of* V. *If* $a, b \in \mathfrak{O}$, *then*

$$a \equiv b \pmod{\mathfrak{a}} \rightarrow V(a) = V(b)$$

Proof. If $a, b \in \mathfrak{a}$, then $0Ra$ and $0Rb$ so that by S1, $0R(a - b)$ and $a - b \in \mathfrak{a}$. If $a \in \mathfrak{a}$ and $b \in \mathfrak{O}$, then $0Ra$ and bRb so that by S2, $0bRab$ therefore, $ab \in \mathfrak{a}$.

Now if $a, b \in \mathfrak{O}$ and if $a - b \in \mathfrak{a}$, then $0R(a - b)$, and since R is transitive,

$$aR0 \,\&\, 0R(a - b) \rightarrow aR(a - b)$$
$$bR0 \,\&\, 0R(a - b) \rightarrow bR(a - b)$$

Then, by S1,

$$aRa \,\&\, aR(a - b) \rightarrow aR(a - (a - b))$$
$$bR(a - b) \,\&\, bR(-b) \rightarrow bR((a - b) - (-b))$$

i.e. aRb and bRa so that $V(a) \leqslant V(b)$ and $V(b) \leqslant V(a)$ and therefore, $V(a) = V(b)$.

We may notice that if $a, b \in \mathfrak{O}$ and if $V(a) = V(b)$, it is not necessarily true that a is congruent to b, modulo $\mathfrak{a}$.

Now if ϕ is a homomorphism of $\mathfrak{O}$ onto a ring $\mathfrak{O}'$, and if the kernel of ϕ is contained in the kernel $\mathfrak{a}$ of V, then, by Proposition 1, one can define a function V' of $\mathfrak{O}'$ onto M by setting $V'(\phi(a)) = V(a)$ for all $a \in \mathfrak{O}$, and it can easily be verified that V' is a valuation of $\mathfrak{O}'$ with M as ordered semi-group of values. We will call V' the projection of V by ϕ. Notice that the kernel of V' is just $\phi(\mathfrak{a})$.

Conversely, if V' is a valuation of $\mathfrak{O}'$ and if for all $a \in \mathfrak{O}$ one sets $V(a) = V'(\phi(a))$, one can easily verify that V is a valuation of $\mathfrak{O}$ and that the kernel of ϕ is contained in the kernel of V.

If $\mathfrak{O}$ is a commutative ring with unity element, then the relation of divisibility "a divides b if and only if there exists an element $c \in \mathfrak{O}$ such that $b = ac$" is evidently a subvaluation of $\mathfrak{O}$. By this definition of divisibility, every element of $\mathfrak{O}$ divides 0 so that we will use the term "regular element" to denote those elements $a \in \mathfrak{O}$ that have the properties "$a \neq 0$ and for all $b \in \mathfrak{O}$, $ab = 0$ implies that $b = 0$," instead of the usual term "non-divisor of zero."

From now on, $\mathfrak{O}$ will always denote a commutative ring with unity element in which every regular element is invertible. Also, G will denote the totally ordered additive group of ordinary integers, G' will denote the totally ordered semi-group obtained by adding the symbol ∞ to G with the laws

1. for all $u \in G'$, $u \leqslant \infty$;
2. for all $u \in G'$, $u + \infty = \infty + u = \infty$;

and G'' will denote the totally ordered semi-group obtained by adding the symbol ∞' to G' with the laws:

1. for all $u \in G''$, $\infty' \leqslant u$;
2. for all $u \in G$, $\infty' + u = u + \infty' = \infty'$;
3. $\infty' + \infty' = \infty'$;
4. $\infty' + \infty = \infty + \infty' = \infty$.

DEFINITION. *We will say that a valuation V of $\mathfrak{O}$ is special if G' is the ordered semi-group of values of V and if there exists a regular element $a \in \mathfrak{O}$ such that $V(a) > 0$.*

If V is a special valuation of $\mathfrak{O}$, there exists an element $a \in \mathfrak{O}$ such that $V(a) \neq \infty$ and then,

$$V(a) = V(1a) = V(1) + V(a)$$

so that $V(1) = 0$. Also, if a is a regular element of $\mathfrak{O}$, then

$$0 = V(1) = V(aa^{-1}) = V(a) + V(\bar{a}^{-1}),$$

so that $V(a) \neq \infty$.

LEMMA 1. *If V is a special valuation of $\mathfrak{O}$ and if we set*

$$\mathfrak{o} = \{a \in \mathfrak{O} | V(a) > 0\}$$

then $\mathfrak{o}$ is an order of $\mathfrak{O}$. We will say that $\mathfrak{o}$ is the order of $\mathfrak{O}$ determined by V.

If $\mathfrak{o}'$ is any order of $\mathfrak{O}$ with the property that for all $a \in \mathfrak{o}'$, $V(a) \geqslant 0$, and if for each positive integer n we set

$$\mathfrak{q}_n = \{a \in \mathfrak{o}' | V(a) \geqslant n\}$$

then $\mathfrak{p} = \mathfrak{q}_1$ is a proper prime ideal of $\mathfrak{o}'$ containing a regular element and the $\mathfrak{q}_n$ are all $\mathfrak{p}$-primary ideals of $\mathfrak{o}'$. Furthermore, if we set

$$\mathfrak{p}' = \{a \in \mathfrak{o}' | V(a) = \infty\}$$

then $\mathfrak{p}'$ is a prime ideal of $\mathfrak{o}'$ and

$$\mathfrak{p}' = \bigcap_{n=1}^{\infty} \mathfrak{q}_n .$$

Finally, if $\mathfrak{o} = \mathfrak{o}'$, the $\mathfrak{q}_n$ are all distinct and $\mathfrak{p}'$ is a prime ideal of $\mathfrak{O}$.

Proof. If $a, b \in \mathfrak{o}$, then $V(a) > 0$ and $V(b) > 0$ so that

$$V(a - b) \geqslant \min\{V(a), V(b)\} > 0, \quad V(ab) = V(a) + V(b) > 0$$

and therefore, $a - b$, $ab \in \mathfrak{o}$. Since $V(1) = 0$, $1 \in \mathfrak{o}$. By the definition of a special valuation, there exists a regular element $a \in \mathfrak{O}$ such that $V(a) > 0$ so that $a \in \mathfrak{o}$ and if b is any element of $\mathfrak{o}$ that is not in $\mathfrak{o}$, then there exists a positive integer n such that $V(a^n) = nV(a) \geqslant -V(b)$ so that

$$V(a^n b) = V(a^n) + V(b) > 0$$

and $a^n b \in \mathfrak{o}$. We have thus shown that $\mathfrak{o}$ is an order of $\mathfrak{O}$.

If $a, b \in \mathfrak{q}_n$, then $V(a) \geqslant n$ and $V(b) \geqslant n$ and therefore,

$$V(a - b) \geqslant \min\{V(a), V(b)\} \geqslant n$$

so that $a - b \in \mathfrak{q}_n$. If $a \in \mathfrak{q}_n$ and $b \in \mathfrak{o}'$, then $V(a) \geqslant n$ and $V(b) \geqslant 0$ so that

$$V(ab) = V(a) + V(b) \geqslant n$$

and therefore, $ab \in \mathfrak{q}_n$. The $\mathfrak{q}_n$ are thus ideals of $\mathfrak{o}'$.

If $a \in \mathfrak{p}$, then $V(a) \geqslant 1$, so that $V(a^n) = nV(a) \geqslant n$ and therefore, $a^n \in \mathfrak{q}_n$. Therefore, $\mathfrak{p} \subseteq \operatorname{rad} \mathfrak{q}_n$. If $a, b \in \mathfrak{o}'$, $a \notin \mathfrak{p}$ and $ab \in \mathfrak{q}_n$, then $V(a) = 0$ and $V(ab) \geqslant n$ so that

$$V(b) = V(a) + V(b) = V(ab) \geqslant n$$

and therefore, $b \in \mathfrak{q}_n$. This proves that $\mathfrak{p}$ is a prime ideal of $\mathfrak{o}'$ and that each $\mathfrak{q}_n$ is $\mathfrak{p}$-primary. Since $V(1) = 0$, $1 \notin \mathfrak{p}$ so that $\mathfrak{p}$ is a proper ideal of $\mathfrak{o}'$. By the definition of a special valuation, there exists a regular element $a \in \mathfrak{o}$ such that $V(a) > 0$, and since $\mathfrak{o}'$ is an order of $\mathfrak{D}$, there exists a regular element $b \in \mathfrak{o}'$ such that $ba = c \in \mathfrak{o}'$. Then c is a regular element of $\mathfrak{o}'$ and $V(c) = V(b) + V(a) > 0$ so that $c \in \mathfrak{p}$.

It is evident that

$$\mathfrak{p}' = \bigcap_{n=1}^{\infty} \mathfrak{q}_n$$

so that $\mathfrak{p}'$ is an ideal of $\mathfrak{o}'$. If a and b are elements of $\mathfrak{o}'$ that are not in $\mathfrak{p}'$, then $V(a) \neq \infty \neq V(b)$ so that $V(ab) = V(a) + V(b) \neq \infty$ and therefore, $ab \notin \mathfrak{p}'$. Thus, $\mathfrak{p}'$ is a prime ideal of $\mathfrak{o}'$.

If $\mathfrak{o} = \mathfrak{o}'$, then evidently,

$$\mathfrak{p}' = \{a \in \mathfrak{D} \mid V(a) = \infty\}.$$

This means that $\mathfrak{p}'$ is the kernel of V so that it is an ideal of $\mathfrak{D}$. Just as above, one may then show that $\mathfrak{p}'$ is a prime ideal of $\mathfrak{D}$. Finally, if n is any positive integer, there exists an element $a \in \mathfrak{D}$ such that $V(a) = n$. This means that $a \in \mathfrak{q}_n$ and $a \notin \mathfrak{q}_{n+1}$. Therefore, the $\mathfrak{q}_n$ are all distinct.

3. The Theory of Quasi-Divisibility. In this section, we give an account of the theory of quasi-divisibility of van der Waerden and Artin, generalized to the case of an arbitrary commutative ring with unity element. The proofs are easy generalizations of those given in **(7, §105)** and will be left to the reader.

Let $\mathfrak{o}$ be an order of $\mathfrak{D}$. By an $\mathfrak{o}$-ideal of $\mathfrak{D}$, we mean simply an $\mathfrak{o}$-submodule of $\mathfrak{D}$, and we denote the set of all $\mathfrak{o}$-ideals of $\mathfrak{D}$ by $\mathfrak{L}(\mathfrak{o})$. The ordinary ideals of $\mathfrak{o}$ are just those $\mathfrak{o}$-ideals of $\mathfrak{D}$ that are contained in $\mathfrak{o}$; we will call these the integral $\mathfrak{o}$-ideals of $\mathfrak{D}$ or just simply the ideals of $\mathfrak{o}$.

If as usual, one defines the product of two $\mathfrak{o}$-ideals $\mathfrak{a}$ and $\mathfrak{b}$ of $\mathfrak{D}$ to be the $\mathfrak{o}$-ideal of $\mathfrak{D}$ generated by the set of all products ab, where $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$,

then with respect to this operation and ordinary set inclusion, $\mathfrak{L}(\mathfrak{o})$ is a commutative partially ordered semi-group with the following properties:

1. $\mathfrak{o}$ is the unity element of $\mathfrak{L}(\mathfrak{o})$;
2. The null ideal (0) is the zero element of $\mathfrak{L}(\mathfrak{o})$;
3. $\mathfrak{L}(\mathfrak{o})$ is complete and distributive, i.e., if $\{\mathfrak{a}_i\}_{i \in I}$ is a set of elements of $\mathfrak{L}(\mathfrak{o})$, then $\Sigma \mathfrak{a}_i$ (the sum of the $\mathfrak{a}_i$ considered as $\mathfrak{o}$-submodules of $\mathfrak{D}$) is the least upper bound of $\{\mathfrak{a}_i\}_{i \in I}$;

$$\bigcap_{i \in I} \mathfrak{a}_i$$

(set-theoretical intersection) is the greatest lower bound of $\{\mathfrak{a}_i\}_{i \in I}$ and if $\mathfrak{a} \in \mathfrak{L}(\mathfrak{o})$, then

$$\mathfrak{a} \sum_{i \in I} \mathfrak{a}_i = \sum_{i \in I} \mathfrak{a}\mathfrak{a}_i.$$

If $\mathfrak{a} \in \mathfrak{L}(\mathfrak{o})$, we denote by $\mathfrak{a}^{-1}$ the set of all $a \in \mathfrak{D}$ such that $a\mathfrak{a} \subseteq \mathfrak{o}$. One may show that $\mathfrak{a}^{-1}$ is the largest $\mathfrak{o}$-ideal $\mathfrak{b}$ of $\mathfrak{D}$ with the property $\mathfrak{b}\mathfrak{a} \subseteq \mathfrak{o}$, and if $\mathfrak{a}$ is invertible, then its inverse is $\mathfrak{a}^{-1}$. If a is a regular element of $\mathfrak{o}$, then $(a)^{-1} = (a^{-1})$.

DEFINITION. *If* $\mathfrak{a}, \mathfrak{b} \in \mathfrak{L}(\mathfrak{o})$, *we say that* $\mathfrak{a}$ *quasi-divides* $\mathfrak{b}$ *and write* $\mathfrak{a} < \mathfrak{b}$, *if* $\mathfrak{a}^{-1} \subseteq \mathfrak{b}^{-1}$.

Let us notice that $\mathfrak{a} \subseteq \mathfrak{b}$ implies that $\mathfrak{a} > \mathfrak{b}$.

The relation of quasi-divisibility is a reflexive and transitive relation on $\mathfrak{L}(\mathfrak{o})$, and furthermore, it is multiplicative. We may then define a relation of "quasi-equality" on $\mathfrak{L}(\mathfrak{o})$ as follows: $\mathfrak{a}$ is quasi-equal to $\mathfrak{b}$, which we write $\mathfrak{a} \sim \mathfrak{b}$, if $\mathfrak{a} < \mathfrak{b}$ and $\mathfrak{b} < \mathfrak{a}$, i.e. if $\mathfrak{a}^{-1} = \mathfrak{b}^{-1}$, and this relation is a multiplicative equivalence relation on $\mathfrak{L}(\mathfrak{o})$. We will denote with $\overline{\mathfrak{L}(\mathfrak{o})}$ the set of equivalence classes of $\mathfrak{L}(\mathfrak{o})$, determined by the relation of quasi-equality, and if $\mathfrak{a} \in \mathfrak{L}(\mathfrak{o})$, we will denote by $\overline{\mathfrak{a}}$ that class in $\overline{\mathfrak{L}(\mathfrak{o})}$ that contains $\mathfrak{a}$. Since the relation of quasi-equality is multiplicative, one can define an induced product on $\overline{\mathfrak{L}(\mathfrak{o})}$ and the relation of quasi-divisibility induces a partial ordering relation on $\overline{\mathfrak{L}(\mathfrak{o})}$, which we denote with the same symbol " $<$ ", and which is multiplicative, so that $\overline{\mathfrak{L}(\mathfrak{o})}$ is a partially ordered semi-group with $\overline{\mathfrak{o}}$ as unity element and $\overline{(0)}$ as zero element.

PROPOSITION 2. *If* $\mathfrak{a}, \mathfrak{b} \in \mathfrak{L}(\mathfrak{o})$, *then* $\mathfrak{a} < \mathfrak{b}$ *if and only if* $(\mathfrak{a}^{-1})^{-1} \supseteq \mathfrak{b}$.

COROLLARY 1. *For all* $\mathfrak{a} \in \mathfrak{L}(\mathfrak{o})$, $\mathfrak{a} \sim (\mathfrak{a}^{-1})^{-1}$ *and for all* $\mathfrak{b} \in \mathfrak{L}(\mathfrak{o})$, $\mathfrak{a} \sim \mathfrak{b}$ *implies that* $(\mathfrak{a}^{-1})^{-1} \supseteq \mathfrak{b}$.

From now on, for all $\mathfrak{a} \in \mathfrak{L}(\mathfrak{o})$, we will denote $(\mathfrak{a}^{-1})^{-1}$ by $\mathfrak{a}^*$. Of course, if $\mathfrak{a}$ is invertible, $\mathfrak{a} = \mathfrak{a}^*$; this is true in particular for the principal $\mathfrak{o}$-ideals of $\mathfrak{D}$ generated by regular elements.

COROLLARY 2. If $\mathfrak{a} \in \mathfrak{L}(\mathfrak{o})$, *then* $\mathfrak{a} > \mathfrak{o}$ *if and only if* $\mathfrak{a} \subseteq \mathfrak{o}$.

COROLLARY 3. *If* $\{\mathfrak{a}_i\}_{i \in I}$ *is a set of elements of* $\mathfrak{L}(\mathfrak{o})$, *then*

$$\overline{\sum_{i \in I} \mathfrak{a}_i} \quad \text{and} \quad \overline{\bigcap_{i \in I} \mathfrak{a}_i^*}$$

are, respectively, the greatest lower bound and least upper bound of $\{\bar{\mathfrak{a}}_i\}_{i \in I}$ *in* $\overline{\mathfrak{L}(\mathfrak{o})}$.

Since the relation of quasi-divisibility is a reflexive, transitive and multiplicative relation on $\mathfrak{L}(\mathfrak{o})$, it is evident that if we define the relation $R_\mathfrak{o}$ on $\mathfrak{D}$ as follows: for all $a, b \in \mathfrak{D}$, $aR_\mathfrak{o}b$ if and only if $(a) \leqslant (b)$, this relation is reflexive, transitive and multiplicative. Then, if $a, b, c \in \mathfrak{D}$ and if $aR_\mathfrak{o}b$ and $aR_\mathfrak{o}c$, then $(a) \leqslant (b)$ and $(a) \leqslant (c)$, which implies that $(a) \leqslant (b) + (c)$. But since

$$b - c \in (b) + (c), \quad (b - c) \geqslant (b) + (c) \geqslant (a)$$

so that $aR_\mathfrak{o}(b - c)$. Therefore, $R_\mathfrak{o}$ is a subvaluation of $\mathfrak{D}$.

PROPOSITION 3. *If* $\mathfrak{a}$, $\mathfrak{b}$ *and* $\mathfrak{c}$ *are ideals of* $\mathfrak{o}$, *if* $\mathfrak{a} + \mathfrak{b} \sim \mathfrak{o}$ *and if* $\mathfrak{a} + \mathfrak{c} \sim \mathfrak{o}$, *then* $\mathfrak{a} + \mathfrak{b}\mathfrak{c} \sim \mathfrak{o}$.

PROPOSITION 4. *If* $\mathfrak{a}$ *and* $\mathfrak{b}$ *are ideals of* $\mathfrak{o}$ *and if* $\mathfrak{a} + \mathfrak{b} \sim \mathfrak{o}$, *then* $\mathfrak{a} \cap \mathfrak{b} \sim \mathfrak{a}\mathfrak{b}$.

PROPOSITION 5. *If* $\mathfrak{a} \in \mathfrak{L}(\mathfrak{o})$ *and if* $\bar{\mathfrak{a}}$ *is invertible in* $\overline{\mathfrak{L}(\mathfrak{o})}$, *then* $\overline{\mathfrak{a}^{-1}}$ *is the inverse of* $\bar{\mathfrak{a}}$.

PROPOSITION 6. *If* $\mathfrak{a}, \mathfrak{b} \in \mathfrak{L}(\mathfrak{o})$, *if* $\bar{\mathfrak{b}}$ *is invertible in* $\overline{\mathfrak{L}(\mathfrak{o})}$ *and if* $\mathfrak{a}:\mathfrak{b}$ *denotes the set of all* $a \in \mathfrak{D}$ *such that* $a\mathfrak{b} \subseteq \mathfrak{a}$, *then* $\mathfrak{a}:\mathfrak{b} \sim \mathfrak{a}\mathfrak{b}^{-1}$.

DEFINITION. *We will say that an* $\mathfrak{o}$*-ideal* $\mathfrak{a}$ *of* $\mathfrak{D}$ *is regular, if*

1. $\mathfrak{a}$ *contains a regular element (of* $\mathfrak{D}$ *or of* $\mathfrak{o}$, *both statements are equivalent),*

2. *there is a regular element* a *(in* $\mathfrak{D}$ *or in* $\mathfrak{o}$, *both statements are equivalent) such that* $a\mathfrak{a} \subseteq \mathfrak{o}$, *i.e.* $\mathfrak{a}^{-1}$ *contains a regular element.*

Let $\mathfrak{F}(\mathfrak{o})$ denote the set of all regular $\mathfrak{o}$-ideals of $\mathfrak{D}$. Then $\mathfrak{F}(\mathfrak{o})$ has the following properties:

1. $\mathfrak{F}(\mathfrak{o})$ is closed under multiplication.
2. $\mathfrak{o} \in \mathfrak{F}(\mathfrak{o})$.
3. If $\mathfrak{a} \in \mathfrak{F}(\mathfrak{o})$, then $\mathfrak{a}^{-1} \subseteq \mathfrak{F}(\mathfrak{o})$.
4. If $\mathfrak{a}, \mathfrak{b} \in \mathfrak{F}(\mathfrak{o})$, then $\mathfrak{a} + \mathfrak{b} \in \mathfrak{F}(\mathfrak{o})$ and $\mathfrak{a} \cap \mathfrak{b} \in \mathfrak{F}(\mathfrak{o})$.
5. If $\{\mathfrak{a}_i\}_{i \in I}$ is a set of elements of $\mathfrak{F}(\mathfrak{o})$ and if this set has an upper bound in $\mathfrak{F}(\mathfrak{o})$, then

$$\sum \mathfrak{a}_i \in \mathfrak{F}(\mathfrak{o}),$$

while if this set has a lower bound in $\mathfrak{F}(\mathfrak{o})$, then

$$\bigcap \mathfrak{a}_i \in \mathfrak{F}(\mathfrak{o}).$$

Since the distributive law is valid on $\mathfrak{F}(\mathfrak{o})$, $\mathfrak{F}(\mathfrak{o})$ is a complete lattice-ordered semi-group (**1**, p. 201).

Let us denote by $\overline{\mathfrak{F}(\mathfrak{o})}$, the set of all $\bar{\mathfrak{a}}$, where $\mathfrak{a} \in \mathfrak{F}(\mathfrak{o})$. We may notice that although $\mathfrak{a} \notin \mathfrak{F}(\mathfrak{o})$, it may very well be that $\mathfrak{a}^{-1} \in \mathfrak{F}(\mathfrak{o})$ and therefore, $\bar{\mathfrak{a}} \in \overline{\mathfrak{F}(\mathfrak{o})}$. It can easily be shown that $\overline{\mathfrak{F}(\mathfrak{o})}$ is also a complete lattice-ordered semi-group.

DEFINITION. An element $a \in \mathfrak{O}$ is said to depend almost integrally on $\mathfrak{o}$ if there exists a regular element $b \in \mathfrak{o}$ such that $ba^n \in \mathfrak{o}$ for all natural integers n. If every element of $\mathfrak{O}$ that depends almost integrally on $\mathfrak{o}$ is in $\mathfrak{o}$, then $\mathfrak{o}$ is said to be *fully integrally closed* in $\mathfrak{O}$.

If $\mathfrak{o}$ is fully integrally closed in $\mathfrak{O}$ and if the ascending chain condition is valid for the ideals of $\mathfrak{o}$, then $\mathfrak{o}$ is integrally closed in $\mathfrak{O}$, that is, every element of $\mathfrak{O}$ which is a root of a monic polynomial with coefficients in $\mathfrak{o}$, is in $\mathfrak{O}$.

THEOREM 1. *If $\mathfrak{o}$ is an order of $\mathfrak{O}$, then $\overline{\mathfrak{F}(\mathfrak{o})}$ is a group if and only if $\mathfrak{o}$ is fully integrally closed in $\mathfrak{O}$.*

4. Definition and elementary Properties of Factorization Rings. Let $\mathfrak{o}$ be an order of $\mathfrak{O}$ which is fully integrally closed in $\mathfrak{O}$ so that $\overline{\mathfrak{F}(\mathfrak{o})}$ is a group. From the theory of partially ordered groups, we know that a unique factorization theorem is valid for the elements of $\overline{\mathfrak{F}(\mathfrak{o})}$ if and only if $\overline{\mathfrak{F}(\mathfrak{o})}$ satisfies the chain condition, that is, if $\bar{\mathfrak{a}}_1 \geq \bar{\mathfrak{a}}_2 \geq \bar{\mathfrak{a}}_3 > \dots$ is a properly descending chain of elements of $\overline{\mathfrak{F}(\mathfrak{o})}$, where $\bar{\mathfrak{a}}_i > \bar{\mathfrak{o}}$ for all indices i, then this chain is finite. When $\overline{\mathfrak{F}(\mathfrak{o})}$ satisfies this condition, then we will say that $\mathfrak{o}$ is a factorization ring. Let us recall that an element $\bar{\mathfrak{p}}$ of $\overline{\mathfrak{F}(\mathfrak{o})}$ is a prime element if and only if $\bar{\mathfrak{p}} > \bar{\mathfrak{o}}$ and $\bar{\mathfrak{p}} > \bar{\mathfrak{a}} > \bar{\mathfrak{o}}$ implies that $\bar{\mathfrak{a}} = \bar{\mathfrak{o}}$ (one may say "for all $\bar{\mathfrak{a}} \in \overline{\mathfrak{F}(\mathfrak{o})}$" or "for all $\bar{\mathfrak{a}} \in \overline{\mathfrak{L}(\mathfrak{o})}$," since $\bar{\mathfrak{p}} > \bar{\mathfrak{a}} > \bar{\mathfrak{o}}$ and $\bar{\mathfrak{p}} \in \overline{\mathfrak{F}(\mathfrak{o})}$ imply that $\bar{\mathfrak{a}} \in \overline{\mathfrak{F}(\mathfrak{o})}$), and that the unique factorization theorem states that if $\bar{\mathfrak{a}} \in \overline{\mathfrak{F}(\mathfrak{o})}$, then

$$\bar{\mathfrak{a}} = \bar{\mathfrak{p}}_1^{m_1}\bar{\mathfrak{p}}_2^{m_2} \dots \bar{\mathfrak{p}}_r^{m_r}$$

where the $\bar{\mathfrak{p}}_i$ are prime elements of $\overline{\mathfrak{F}(\mathfrak{o})}$, this decomposition being unique, and $\bar{\mathfrak{a}} > \bar{\mathfrak{o}}$ if and only if $n_i > 0$ for all indices i.

Throughout this section, we assume that $\mathfrak{o}$ is a factorization ring with $\mathfrak{O}$ as full ring of quotients.

THEOREM 2. *If $\bar{\mathfrak{p}}$ is a prime element of $\overline{\mathfrak{F}(\mathfrak{o})}$ and if $\mathfrak{p} = \mathfrak{p}^*$, then $\mathfrak{p}$ is a regular prime ideal of $\mathfrak{o}$. We will call these prime ideals the relevant prime ideals of $\mathfrak{o}$.*

Proof. If $a, b \in \mathfrak{o}$, $ab \in \mathfrak{p}$ and $b \notin \mathfrak{p}$, then $(b) + \mathfrak{p} \supset \mathfrak{p} = \mathfrak{p}^*$ so that $\bar{\mathfrak{p}} > \overline{(b) + \mathfrak{p}} > \bar{\mathfrak{o}}$ and therefore, $(b) + \mathfrak{p} \sim \mathfrak{o}$. Then, $(a) = a\mathfrak{o} \sim a((b) + \mathfrak{p}) = (ab) + a\mathfrak{p} \subseteq \mathfrak{p}$ so that $(a) > \mathfrak{p}$ and therefore, $a \in \mathfrak{p}^* = \mathfrak{p}$. Since $\bar{\mathfrak{p}} \in \overline{\mathfrak{F}(\mathfrak{o})}$ and $\mathfrak{p} = \mathfrak{p}^*$, $\mathfrak{p}$ must contain a regular element of $\mathfrak{o}$.

LEMMA 2. *If $\mathfrak{p}$ is a relevant prime ideal of $\mathfrak{o}$ and if $\mathfrak{a}$ and $\mathfrak{b}$ are ideals of $\mathfrak{o}$, then,*

1. *if* $\mathfrak{a} \geqslant \mathfrak{p}^m$, $\mathfrak{a} \not\geqslant \mathfrak{p}^{m+1}$, $\mathfrak{b} \geqslant \mathfrak{p}^n$ *and* $\mathfrak{b} \not\geqslant \mathfrak{p}^{n+1}$, *where* m *and* n *are non-negative integers, then* $\mathfrak{ab} \geqslant \mathfrak{p}^{m+n}$ *and* $\mathfrak{ab} \not\geqslant \mathfrak{p}^{m+n+1}$.

2. *if for all natural numbers* n, $\mathfrak{a} \geqslant \mathfrak{p}^n$, *then for all* n, $\mathfrak{ab} \geqslant \mathfrak{p}^n$.

Proof. First of all, let us suppose that $\mathfrak{a}$ and $\mathfrak{b}$ satisfy the hypothesis of the first case. Then surely, $\mathfrak{ab} \geqslant \mathfrak{p}^{m+n}$. If $\mathfrak{ab} \geqslant \mathfrak{p}^{m+n+1}$, then $(\mathfrak{a}\mathfrak{p}^{-m})\ (\mathfrak{b}\mathfrak{p}^{-n}) \geqslant \mathfrak{p}$ so that $(\mathfrak{a}\mathfrak{p}^{-m})\ (\mathfrak{b}\mathfrak{p}^{-n}) \subseteq \mathfrak{p}$. But since $\mathfrak{a}\mathfrak{p}^{-m} \geqslant \mathfrak{o}$ and $\mathfrak{b}\mathfrak{p}^{-n} \geqslant \mathfrak{o}$, $\mathfrak{a}\mathfrak{p}^{-m}$ and $\mathfrak{b}\mathfrak{p}^{-n}$ are ideals of $\mathfrak{o}$ so that $\mathfrak{a}\mathfrak{p}^{-m} \subseteq \mathfrak{p}$ or $\mathfrak{b}\mathfrak{p}^{-n} \subseteq \mathfrak{p}$ which implies that $\mathfrak{a} \geqslant \mathfrak{p}^{m+1}$ or $\mathfrak{b} \geqslant \mathfrak{p}^{n+1}$, contradicting our hypotheses.

The second case is trivial since $\mathfrak{ab} \subseteq \mathfrak{a}$ and therefore, $\mathfrak{ab} \geqslant \mathfrak{a}$.

Now, if $\mathfrak{p}$ is a relevant prime ideal of $\mathfrak{o}$, we may use $\mathfrak{p}$ to define a valuation V on $\mathfrak{o}$ as follows: if $a \in \mathfrak{o}$ and if there is a non-negative integer n such that $(a) \geqslant \mathfrak{p}^n$ and $(a) \not\geqslant \mathfrak{p}^{n+1}$, we set $V(a) = n$, while if for all natural numbers n, $(a) \geqslant \mathfrak{p}^n$, we set $V(a) = \infty$. By Lemma 2, for all $a, b \in \mathfrak{o}$, $V(ab) = V(a) + V(b)$. Now, if $a, b \in \mathfrak{o}$ and if n is a non-negative integer such that $(a) \geqslant \mathfrak{p}^n$ and $(b) \geqslant \mathfrak{p}^n$, then since $(a - b) \subseteq (a) + (b)$, we have

$$\overline{(a-b)} \geqslant \overline{(a) + (b)} \geqslant \bar{\mathfrak{p}}^n$$

(Cor. 3, Prop. 2) so that $V(a-b) \geqslant \min\{V(a), V(b)\}$. If a is a regular element of $\mathfrak{o}$, since $(a) \in \mathfrak{F}(\mathfrak{o})$, it is evident that $V(a) \neq \infty$. We may then extend the function V to all of $\mathfrak{O}$ as follows: if $a \in \mathfrak{O}$, $a = b/c$, where $b, c \in \mathfrak{o}$ and c is regular, and we set $V(a) = V(b) - V(c)$. Then, one can easily verify that V is well defined on all of $\mathfrak{O}$ and that it satisfies V1′ and V2. If n is any non-negative integer, since $(\mathfrak{p}^{n+1})^* \subset (\mathfrak{p}^n)^*$, there is an element $a \in (\mathfrak{p}^n)^*$ such that $a \notin (\mathfrak{p}^{n+1})^*$ and therefore, $V(a) = n$. Evidently, $V(0) = \infty$. Now $\mathfrak{p}$ contains a regular element a, and therefore, $V(a) = m > 0$, and if n is any positive integer, one can find an integer k such that $V(a^k) = km \geqslant n$ and therefore, there exists an element $b \in \mathfrak{o}$ such that $V(b) = km - n$ so that

$$V(ba^{-k}) = V(b) - V(a^k) = -n.$$

Thus V is a special valuation of $\mathfrak{O}$. We will say that V is the valuation of $\mathfrak{O}$ determined by the relevant prime ideal $\mathfrak{p}$ of $\mathfrak{o}$.

THEOREM 3. *If* $\mathfrak{p}$ *is a relevant prime ideal of* $\mathfrak{o}$ *and if* V *is the valuation of* $\mathfrak{O}$ *determined by* $\mathfrak{p}$, *then for each positive integer* n,

$$\mathfrak{p}^{(n)} = (\mathfrak{p}^n)^* = \{a \in \mathfrak{o} \mid V(a) \geqslant n\}.$$

Proof. If $a \in \mathfrak{o}$, $V(a) \geqslant n$ if and only if $(a) \geqslant \mathfrak{p}^n$, which evidently means that $a \in (\mathfrak{p}^n)^*$, so that

$$(\mathfrak{p}^n)^* = \{a \in \mathfrak{o} \mid V(a) \geqslant n\}.$$

Then, by Lemma 1, each $(\mathfrak{p}^n)^*$ is a $\mathfrak{p}$-primary ideal of $\mathfrak{o}$ and since $\mathfrak{p}^n \subseteq (\mathfrak{p}^n)^*$, $\mathfrak{p}^{(n)} \subseteq (\mathfrak{p}^n)^*$. If $a \in (\mathfrak{p}^n)^*$, then $(a) \geqslant \mathfrak{p}^n$ and therefore, $a(\mathfrak{p}^n)^{-1} \subseteq \mathfrak{o}$. Since $(\mathfrak{p}^n)^{-1}\mathfrak{p}^n \sim \mathfrak{o}$, there exists $b \in (\mathfrak{p}^n)^{-1}\mathfrak{p}^n$ such that $b \notin \mathfrak{p}$. Then

$$ab \in a((\mathfrak{p}^n)^{-1}\mathfrak{p}^n) = (a(\mathfrak{p}^n)^{-1})\mathfrak{p}^n \subseteq \mathfrak{p}^n$$

so that $a \in (\mathfrak{p}^n)_\mathfrak{p} = \mathfrak{p}^{(n)}$ and therefore, $\mathfrak{p}^{(n)} = (\mathfrak{p}^n)^*$.

Under the hypotheses of Theorem 3, if we set

$$\mathfrak{p}' = \{a \in \mathfrak{o} | V(a) = \infty\}$$

by Lemma 1, $\mathfrak{p}'$ is a prime ideal of $\mathfrak{o}$ and

$$\mathfrak{p}' = \bigcap_{n=1}^{\infty} \mathfrak{p}^{(n)} .$$

We will call $\mathfrak{p}'$ the associate of $\mathfrak{p}$.

THEOREM 4. *If $\mathfrak{p}$ is a relevant prime ideal of $\mathfrak{o}$ and if $\mathfrak{p}'$ is the associate of $\mathfrak{p}$, then $\mathfrak{p}'$ has the following properties:*

1. *for every natural integer n, $\mathfrak{p}' > \mathfrak{p}^n$ and $\mathfrak{p}'$ contains any $\mathfrak{o}$-ideal of $\mathfrak{D}$ that has this property,*
2. $(\mathfrak{p})'^* = \mathfrak{p}'$.
3. *any prime ideal of $\mathfrak{o}$ that is properly contained in $\mathfrak{p}$, is contained in $\mathfrak{p}'$.*

Proof. If n is any natural integer, for all $a \in \mathfrak{p}'$, $(a) > \mathfrak{p}^n$, so that by Corollary 3 of Proposition 2,

$$\overline{\mathfrak{p}'} = \overline{\sum_{a \in \mathfrak{p}'} (a)} > \overline{\mathfrak{p}}^n .$$

Then, if $\mathfrak{a}$ is any $\mathfrak{o}$-ideal of $\mathfrak{D}$ with the property that for all natural integers n, $\mathfrak{a} > \mathfrak{p}^n$, for all $a \in \mathfrak{a}$, $(a) > \mathfrak{a} > \mathfrak{p}^n$ so that $V(a) = \infty$ and therefore, $a \in \mathfrak{p}'$ and $\mathfrak{a} \subseteq \mathfrak{p}'$. In particular, $(\mathfrak{p}')^*$ has this property, so that $(\mathfrak{p}')^* = \mathfrak{p}'$.

Let $\mathfrak{a}$ be an ideal of $\mathfrak{o}$ such that $\mathfrak{a} \subset \mathfrak{p}$ and $\mathfrak{a} \not\subseteq \mathfrak{p}'$. Then, there is a non-negative integer n such that $\mathfrak{a} > \mathfrak{p}^n$ and $\mathfrak{a} \not> \mathfrak{p}^{n+1}$. Now $(\mathfrak{a}\mathfrak{p}^{-1})\mathfrak{p} = \mathfrak{a}(\mathfrak{p}^{-1}\mathfrak{p}) \subseteq \mathfrak{a}$. Since $\mathfrak{a} \subset \mathfrak{p}$, $\mathfrak{p} \not\subseteq \mathfrak{a}$ and $\mathfrak{a}\mathfrak{p}^{-1} > \mathfrak{o}$ so that $\mathfrak{a}\mathfrak{p}^{-1} \subseteq \mathfrak{o}$. If $\mathfrak{a}\mathfrak{p}^{-1} \subseteq \mathfrak{a}$, then $\mathfrak{a}\mathfrak{p}^{-1} > \mathfrak{a} > \mathfrak{p}^n$ so that $\mathfrak{a} > \mathfrak{p}^{n+1}$, a contradiction. Therefore, $\mathfrak{a}$ is not a prime ideal of $\mathfrak{o}$.

THEOREM 5. *If $\mathfrak{a}$ is an ideal of $\mathfrak{o}$ such that $\overline{\mathfrak{a}} \in \overline{\mathfrak{F}(\mathfrak{o})}$, and if*

$$\overline{\mathfrak{a}} = \overline{\mathfrak{p}_1}^{n_1}\overline{\mathfrak{p}_2}^{n_2} \cdots \overline{\mathfrak{p}_r}^{n_r}$$

where $\mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_r$ are relevant prime ideals of $\mathfrak{o}$, then

$$\mathfrak{a}^* = \mathfrak{p}_1^{(n_1)} \cap \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)} .$$

Proof. Since $\overline{\mathfrak{a}}$ is the least upper bound, in $\overline{\mathfrak{F}(\mathfrak{o})}$, and therefore also in $\overline{\mathfrak{L}(\mathfrak{o})}$, of the set

$$\{\overline{\mathfrak{p}_1^{n_1}}, \overline{\mathfrak{p}_2^{n_2}}, \ldots, \overline{\mathfrak{p}_r^{n_r}}\}$$

by Corollary 3 of Proposition 2,

$$\mathfrak{a} \sim \mathfrak{p}_1^{(n_1)} \cap \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)}$$

and therefore,

$$\mathfrak{a}^* \supseteq \mathfrak{p}_1^{(n_1)} \cap \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)}$$

But since, for each $i = 1,2, \ldots, r$,

$$\mathfrak{a}^* \sim \mathfrak{a} \geqslant \mathfrak{p}_i^{n_i}, \mathfrak{a}^* \subseteq \mathfrak{p}_i^{(n_i)}$$

and therefore,

$$\mathfrak{a}^* \subseteq \mathfrak{p}_1^{(n_1)} \cap \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)}.$$

COROLLARY 1. *If a is a regular element of $\mathfrak{o}$ and if a is not a unit of $\mathfrak{o}$, then*

$$(a) = \mathfrak{p}_1^{(n_1)} \cap \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)}$$

where $\mathfrak{p}_1,\mathfrak{p}_2, \ldots, \mathfrak{p}_r$ are relevant prime ideals of $\mathfrak{o}$.

COROLLARY 2. *The relevant prime ideals of $\mathfrak{o}$ are just the minimal proper regular prime ideals of $\mathfrak{o}$. Every regular prime ideal of $\mathfrak{o}$ contains a relevant prime ideal of $\mathfrak{o}$ so that the non-minimal regular prime ideals of $\mathfrak{o}$ are all quasi-equal to $\mathfrak{o}$.*

Proof. Since the relevant prime ideals of $\mathfrak{o}$ are regular and since no relevant prime ideal is properly contained in another, all we have to show is that any regular prime ideal $\mathfrak{p}$ of $\mathfrak{o}$ contains a relevant prime ideal of $\mathfrak{o}$.

We need only consider the case where $\mathfrak{p} \neq \mathfrak{o}$, and then, $\mathfrak{p}$ contains a regular element a which is not a unit of $\mathfrak{o}$, so that by Corollary 1 of Theorem 5,

$$(a) = \mathfrak{p}_1^{(n_1)} \cap \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)}$$

where $\mathfrak{p}_1,\mathfrak{p}_2, \ldots, \mathfrak{p}_r$ are relevant prime ideals of $\mathfrak{o}$, and evidently, $\mathfrak{p}$ must contain one of them.

THEOREM 6. *If $\mathfrak{p}$ is a relevant prime ideal of $\mathfrak{o}$, then the only $\mathfrak{p}$-primary ideals of $\mathfrak{o}$ are the formal powers of $\mathfrak{p}$.*

Proof. Let $\mathfrak{q}$ be a $\mathfrak{p}$-primary ideal of $\mathfrak{o}$. Since $\mathfrak{p}$ contains a regular element a and since $\operatorname{rad} \mathfrak{q} = \mathfrak{p}$, $a^m \in \mathfrak{q}$ for some positive integer m and $\mathfrak{q}$ is regular. Since a^m is not a unit of $\mathfrak{o}$, by Corollary 1 of Theorem 5,

$$(a^m) = \mathfrak{p}_1^{(n_1)} \cap \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)}$$

where $\mathfrak{p}_1,\mathfrak{p}_2, \ldots, \mathfrak{p}_r$ are relevant prime ideals of $\mathfrak{o}$, and $\mathfrak{p}$ must contain one of these, say $\mathfrak{p} \supseteq \mathfrak{p}_1$. Since $\mathfrak{p}$ is relevant, $\mathfrak{p} = \mathfrak{p}_1$. Then $\mathfrak{p}^{(n_1)} = (a^m)_{\mathfrak{p}}$ and since $\mathfrak{q}$ is a $\mathfrak{p}$-primary ideal containing (a^m),

$$\mathfrak{q} \supseteq \mathfrak{p}^{(n_1)}$$

so that

$$\mathfrak{q} \leqslant \mathfrak{p}^{n_1}$$

and therefore, $\mathfrak{q} \sim \mathfrak{p}^n$ for some positive integer $n \leqslant n_1$ and $\mathfrak{q} \subseteq \mathfrak{p}^{(n)}$. Then, by Proposition 6,

$$\mathfrak{q}:\mathfrak{p}^{(n)} \sim \mathfrak{q}(\mathfrak{p}^{(n)})^{-1} \sim \mathfrak{q}(\mathfrak{p}^n)^{-1} \sim \mathfrak{o}$$

so that there exists an element $b \in \mathfrak{q}:\mathfrak{p}^{(n)}$ such that $b \notin \mathfrak{p}$. Then, $b\mathfrak{p}^{(n)} \subseteq \mathfrak{q}$, and since $b \notin \mathfrak{p}$ and $\mathfrak{q}$ is $\mathfrak{p}$-primary, $\mathfrak{p}^{(n)} \subseteq \mathfrak{q}$.

COROLLARY. *If $\mathfrak{a}$ is an ideal of $\mathfrak{o}$ and if*

$$\infty \neq n = \min \{V(a) | a \in \mathfrak{a}\}$$

then $\mathfrak{a}_\mathfrak{p} = \mathfrak{p}^{(n)}$ (by definition, $\mathfrak{p}^{(0)} = \mathfrak{o}$).

Proof. It is evident that $\mathfrak{a} \subseteq \mathfrak{p}^{(n)}$ and that $\mathfrak{a} \not\subseteq \mathfrak{p}^{(n+1)}$. If $\mathfrak{p}$ were not a minimal prime ideal of $\mathfrak{a}$, then $\mathfrak{p}$ would properly contain a minimal prime ideal of $\mathfrak{a}$, which, by Theorem 4, would be contained in $\mathfrak{p}'$ so that $\mathfrak{a} \subseteq \mathfrak{p}'$, a contradiction. Therefore, $\mathfrak{p}$ is a minimal prime ideal of $\mathfrak{a}$ and $\mathfrak{a}_\mathfrak{p}$ is $\mathfrak{p}$-primary. Then, since $\mathfrak{a} \subseteq \mathfrak{p}^{(n)}$, $\mathfrak{a}_\mathfrak{p} \subseteq \mathfrak{p}^{(n)}$ and since $\mathfrak{a} \subseteq \mathfrak{a}_\mathfrak{p}$, $\mathfrak{a}_\mathfrak{p} \not\subseteq \mathfrak{p}^{(n+1)}$ so that by Theorem 6, $\mathfrak{a}_\mathfrak{p} = \mathfrak{p}^{(n)}$.

THEOREM 7. *If $\mathfrak{p}$ is a relevant prime ideal of $\mathfrak{o}$, if $\mathfrak{p}'$ is the associate of $\mathfrak{p}$ and if V is the valuation of $\mathfrak{D}$ determined by $\mathfrak{p}$, then $(\mathfrak{o}/\mathfrak{p}')_{(\mathfrak{p}/\mathfrak{p}')}$ is a regular local ring of dimension 1 (valuation ring determined by a discrete, rank 1 valuation of its field of quotients) and if ϕ denotes the natural homomorphism of $\mathfrak{o}$ onto $\mathfrak{o}/\mathfrak{p}'$, then, for any two elements $a, b \in \mathfrak{o}$, $V(a) \leqslant V(b)$ if and only if $\phi(a)$ divides $\phi(b)$ in $(\mathfrak{o}/\mathfrak{p}')_{(\mathfrak{p}/\mathfrak{p}')}$.*

Proof. Since $\mathfrak{p}'$ is a prime ideal of $\mathfrak{o}$, $\mathfrak{o}/\mathfrak{p}'$ and $(\mathfrak{o}/\mathfrak{p}')_{(\mathfrak{p}/\mathfrak{p}')}$ are integral domains. Since every prime ideal of $\mathfrak{o}$ that is properly contained in $\mathfrak{p}$ must be contained in $\mathfrak{p}'$ (Theorem 4), the null ideal is the only prime ideal of $\mathfrak{o}/\mathfrak{p}'$ that is properly contained in $\mathfrak{p}/\mathfrak{p}'$ so that $(\mathfrak{o}/\mathfrak{p}')_{(\mathfrak{p}/\mathfrak{p}')}$ contains only one non-trivial prime ideal. Then, an arbitrary non-trivial ideal of $(\mathfrak{o}/\mathfrak{p}')_{(\mathfrak{p}/\mathfrak{p}')}$ must be a primary ideal belonging to this unique non-trivial prime ideal. Now, the only $\mathfrak{p}$-primary ideals of $\mathfrak{o}$ are the formal powers of $\mathfrak{p}$ (Theorem 6) and they are totally ordered under ordinary inclusion and all contain $\mathfrak{p}'$ so that the $\mathfrak{p}/\mathfrak{p}'$-primary ideals of $\mathfrak{o}/\mathfrak{p}'$ are totally ordered under ordinary inclusion and therefore, the non-trivial ideals of $(\mathfrak{o}/\mathfrak{p}')_{(\mathfrak{p}/\mathfrak{p}')}$ are also totally ordered under ordinary inclusion. Thus, $(\mathfrak{o}/\mathfrak{p}')_{(\mathfrak{p}/\mathfrak{p}')}$ is a regular local ring of dimension 1.

If $a, b \in \mathfrak{o}$, $V(a) \leqslant V(b)$ if and only if for all non-negative integers n, $(a) \geqslant \mathfrak{p}^n$ implies that $(b) \geqslant \mathfrak{p}^n$, that is, $a \in \mathfrak{p}^{(n)}$ implies that $b \in \mathfrak{p}^{(n)}$ or $\phi(a) \in \phi(\mathfrak{p}^{(n)})$ implies that $\phi(b) \in \phi(\mathfrak{p}^{(n)})$, which evidently means that $\phi(a)$ divides $\phi(b)$ in $(\mathfrak{o}/\mathfrak{p}')_{(\mathfrak{p}/\mathfrak{p}')}$.

Remarks. If $\mathfrak{o}$ is an order of $\mathfrak{D}$, if $\mathfrak{o}$ is fully integrally closed in $\mathfrak{D}$ and if the ascending chain condition holds for the ideals of $\mathfrak{o}$, then $\mathfrak{o}$ is surely a factorization ring, for if $\overline{\mathfrak{a}_1} > \overline{\mathfrak{a}_2} > \overline{\mathfrak{a}_3} > \ldots$ is a chain of elements of $\overline{\mathfrak{F}(\mathfrak{o})}$ and if, for each index i, $\overline{\mathfrak{a}_i} > \overline{\mathfrak{o}}$, then the ascending chain $\mathfrak{a}_1^* \subset \mathfrak{a}_2^* \subset \mathfrak{a}_3^* \subset \ldots$ of ideals of $\mathfrak{o}$ must be finite. Furthermore, by Krull's Intersection Theorem,

$$\mathfrak{p}' = \bigcap_{n=1}^{\infty} \mathfrak{p}^{(n)} = (0)_{\mathfrak{p}}$$

so that $(\mathfrak{o}/\mathfrak{p}')_{(\mathfrak{p}/\mathfrak{p}')}$ is just the generalized ring of quotients $\mathfrak{o}_{\mathfrak{p}}$ and $\mathfrak{p}'$ is a minimal prime ideal of $\mathfrak{o}$. (We have thus obtained the results of (**5**, §4.7 and §4.9).)

Under these assumptions, since the associates of relevant prime ideals of $\mathfrak{o}$ are minimal prime ideals of $\mathfrak{o}$, there is only a finite number of them and by Theorem 4, for each relevant prime ideal $\mathfrak{p}$ of $\mathfrak{o}$, $\mathfrak{p}'$ is the only prime ideal of $\mathfrak{o}$ properly contained in $\mathfrak{p}$. If $\mathfrak{a}$ is an ideal of $\mathfrak{o}$ contained in $\mathfrak{p}'$, then

$$\mathfrak{p}' = (0)_{\mathfrak{p}} \subseteq \mathfrak{a}_{\mathfrak{p}} \subseteq \mathfrak{p}_{\mathfrak{p}}' = \mathfrak{p}'$$

so that $\mathfrak{a}_{\mathfrak{p}} = \mathfrak{p}'$. Therefore, $\mathfrak{p}'$ is the only $\mathfrak{p}'$-primary ideal of $\mathfrak{o}$, and if we set

$$\mathfrak{p}' = \bigcap_{n=1}^{\infty} \mathfrak{p}^{(n)} = \mathfrak{p}^{(\infty)}$$

then we may restate the Corollary of Theorem 5 more generally as follows: If $\mathfrak{a}$ is an ideal of $\mathfrak{o}$ and if $n = \min \{V(a) | a \in \mathfrak{a}\}$ (n may equal ∞), then $\mathfrak{a}_{\mathfrak{p}} = \mathfrak{p}^{(n)}$.

THEOREM 8. *If $\{\mathfrak{p}_i\}_{i \in I}$ is the set of relevant prime ideals of $\mathfrak{o}$ and if for each $i \in I$, V_i denotes the valuation of $\mathfrak{D}$ determined by $\mathfrak{p}_i$, then for all $a \in \mathfrak{D}$, $a \in \mathfrak{o}$ if and only if $V_i(a) \geqslant 0$ for all $i \in I$. Furthermore, if for each $i \in I$, R_i is the subvaluation of $\mathfrak{D}$ determined by V_i, then*

$$R_{\mathfrak{o}} = \bigcap_{i \in I} R_i$$

Proof. From the definition of V_i, it is evident that if $a \in \mathfrak{o}$, then $V_i(a) \geqslant 0$ for all $i \in I$. Conversely, let a be an element of $\mathfrak{D}$ such that $V_i(a) \geqslant 0$ for all $i \in I$. Then, $a = b/c$, where $b, c \in \mathfrak{o}$, c is regular and $V_i(b) \geqslant V_i(c)$ for all $i \in I$. If c is not a unit of $\mathfrak{o}$, by Corollary 1 of Theorem 5,

$$(c) = \mathfrak{p}_{i_1}^{(n_1)} \cap \mathfrak{p}_{i_2}^{(n_2)} \cap \ldots \cap \mathfrak{p}_{i_r}^{(n_r)}$$

where $i_1, i_2, \ldots, i_r \in I$. Then, for each $j = 1,2, \ldots, r$,

$$\text{since } V_{i_j}(b) \geqslant V_{i_j}(c),\ (b) \geqslant \mathfrak{p}_{i_j}^{n_j} \text{ so that } b \in \mathfrak{p}_{i_j}^{(n_j)}$$

and therefore, $b \in (c)$ and $a = b/c \in \mathfrak{o}$. We may notice here that for every regular element $c \in \mathfrak{D}$, $V_i(c) = 0$ for all $i \in I$ except a finite number.

If $a, b \in \mathfrak{o}$ and $aR_{\mathfrak{o}}b$, then $(a) \leqslant (b)$ and from the definition of V_i, it is evident that $V_i(a) \leqslant V_i(b)$. In general, if $a, b \in \mathfrak{D}$, $a = c_1/d_1$, $b = c_2/d_2$, where $c_1, d_1, c_2, d_2 \in \mathfrak{o}$ and d_1 and d_2 are regular, and if $(a) \leqslant (b)$, then $(c_1)(d_1)^{-1} \leqslant (c_2)(d_2)^{-1}$ so that $(c_1d_2) \leqslant (c_2d_1)$ and therefore,

$$V_i(c_1) + V_i(d_2) = V_i(c_1d_2) \leqslant V_i(c_2d_1) = V_i(c_2) + V_i(d_1),$$

and this implies that

$$V_i(a) = V_i(c_1) - V_i(d_1) \leqslant V_i(c_2) - V_i(d_2) = V_i(b),$$

i.e. aR_ib. Therefore,

$$R_o \subseteq \bigcap_{i \in I} R_i .$$

Conversely, if $a, b \in \mathfrak{O}$ and if for all $i \in I$, aR_ib, i.e. $V_i(a) \leqslant V_i(b)$, then, if $c \in (a)^{-1}$, $ca \in \mathfrak{o}$ so that for all $i \in I$,

$$V_i(cb) = V_i(c) + V_i(b) \geqslant V_i(c) + V_i(a) = V_i(ca) \geqslant 0$$

and therefore, $cb \in \mathfrak{o}$. Therefore, $(a)^{-1} \subseteq (b)^{-1}$, i.e., $(a) \leqslant (b)$ or aR_ob.

As a particular case of this theorem, we have that if there is only one relevant prime ideal $\mathfrak{p}$ in $\mathfrak{o}$ and if V is the valuation of $\mathfrak{O}$ determined by $\mathfrak{p}$, then $\mathfrak{o}$ is the order of $\mathfrak{O}$ determined by V and R_o is the sub-valuation of $\mathfrak{O}$ determined by V.

5. Two characterizations of factorization rings. In this section, we will see that the properties of factorization rings given by Corollary 1 of Theorem 5 and by Theorem 8 may be used to characterize factorization rings.

If $\mathfrak{o}$ is an order of $\mathfrak{O}$ and if V is a special non-negative valuation of $\mathfrak{O}$, we define the function $V_\mathfrak{o}$ on $\mathfrak{L}(\mathfrak{o})$ as follows: if $\mathfrak{a} \in \mathfrak{L}(\mathfrak{o})$ and if $\{V(a) | a \in \mathfrak{a}\}$ has a minimum, then $V_\mathfrak{o}(\mathfrak{a}) = \min \{V(a) | a \in \mathfrak{a}\}$, while otherwise, $V_\mathfrak{o}(\mathfrak{a}) = \infty'$. It is evident that for all $a \in \mathfrak{O}$, $V_\mathfrak{o}((a)) = V(a)$ so that we may think of $V_\mathfrak{o}$ as an extension of V and drop the subscript $\mathfrak{o}$ where no ambiguity arises.

LEMMA 3. *If $\mathfrak{o}$ is an order of $\mathfrak{O}$ and if V is a special valuation of $\mathfrak{O}$ such that $V(a) \geqslant 0$ for all $a \in \mathfrak{o}$, then $V_\mathfrak{o}$ is a homomorphism of $\mathfrak{L}(\mathfrak{o})$ onto G'', and for all $\mathfrak{a}, \mathfrak{b} \in \mathfrak{L}(\mathfrak{o})$, $V(\mathfrak{a} + \mathfrak{b}) = \min \{V(\mathfrak{a}), V(\mathfrak{b})\}$.*

Proof. Let $\mathfrak{a}, \mathfrak{b} \in \mathfrak{L}(\mathfrak{o})$. It is evident that $\mathfrak{a} \subseteq \mathfrak{b}$ implies that $V(\mathfrak{a}) \geqslant V(\mathfrak{b})$. Any element of $\mathfrak{ab}$ has the form $a_1b_1 + \ldots + a_nb_n$ where $a_i \in \mathfrak{a}$ and $b_i \in \mathfrak{b}$, and

$$\begin{aligned} V(a_1b_1 + \ldots + a_nb_n) &\geqslant \min \{V(a_1) + V(b_1), \ldots, V(a_n) + V(b_n)\} \\ &\geqslant V(\mathfrak{a}) + V(\mathfrak{b}) \end{aligned}$$

so that $V(\mathfrak{ab}) \geqslant V(\mathfrak{a}) + V(\mathfrak{b})$. To establish the reverse inequality, let us suppose first of all that $V(\mathfrak{a}) \neq \infty' \neq V(\mathfrak{b})$ so that there exist $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$ such that $V(\mathfrak{a}) = V(a)$ and $V(\mathfrak{b}) = V(b)$, and then,

$$V(\mathfrak{a}) + V(\mathfrak{b}) = V(a) + V(b) = V(ab) \geqslant V(\mathfrak{ab}).$$

Secondly, let us suppose that $V(\mathfrak{a}) = \infty'$ and that $V(\mathfrak{b}) \neq \infty$. Then, there exists $b \in \mathfrak{b}$ such that $V(b) \neq \infty$ and if n is any ordinary integer, there exists $a \in \mathfrak{a}$ such that $V(a) \leqslant n - V(b)$ so that $V(ab) = V(a) + V(b) \leqslant n$ and therefore, $V(\mathfrak{ab}) = \infty' = V(\mathfrak{a}) + V(\mathfrak{b})$. Finally, if $V(\mathfrak{a}) = \infty'$ and $V(\mathfrak{b}) = \infty$, for every $b \in \mathfrak{b}$, $V(b) = \infty$ so that for every $a \in \mathfrak{a}$, $V(ab) = V(a) + V(b) = \infty$ and therefore, $V(\mathfrak{ab}) = \infty = V(\mathfrak{a}) + V(\mathfrak{b})$. That $V_\mathfrak{o}$ is a function of

$\mathfrak{L}(\mathfrak{o})$ onto G'' is evident, since for all $a \in \mathfrak{D}$, $V((a)) = V(a)$ and since $V(\mathfrak{D}) = \infty'$. If $\mathfrak{a}, \mathfrak{b} \in \mathfrak{L}(\mathfrak{o})$, since $\mathfrak{a} \subseteq \mathfrak{a} + \mathfrak{b}$, $V(\mathfrak{a}) \geq V(\mathfrak{a} + \mathfrak{b})$ and similarly, $V(\mathfrak{b}) \geq V(\mathfrak{a} + \mathfrak{b})$ so that $\min \{V(\mathfrak{a}), V(\mathfrak{b})\} \geq V(\mathfrak{a} + \mathfrak{b})$. Then, if $a \in \mathfrak{a}$ and $b \in \mathfrak{b}$,

$$V(a + b) \geq \min \{V(a), V(b)\} \geq \min \{V(\mathfrak{a}), V(\mathfrak{b})\}$$

so that $V(\mathfrak{a} + \mathfrak{b}) \geq \min \{V(\mathfrak{a}), V(\mathfrak{b})\}$.

LEMMA 4. *If $\mathfrak{o}$ is a factorization ring with $\mathfrak{D}$ as full ring of quotients, if W is a special valuation of $\mathfrak{D}$ such that for all $a \in \mathfrak{o}$, $W(a) \geq 0$, and if $\{a \in \mathfrak{o} | W(a) > 0\}$ is a relevant prime ideal $\mathfrak{p}$ of $\mathfrak{o}$, then W coincides with the valuation V of $\mathfrak{D}$ determined by $\mathfrak{p}$.*

Proof. By Lemma 1, for each positive integer n, the set

$$\mathfrak{q}_n = \{a \in \mathfrak{o} | W(a) \geq n\}$$

is a $\mathfrak{p}$-primary ideal of $\mathfrak{o}$. Since $\mathfrak{p}^n \subseteq \mathfrak{q}_n$, $\mathfrak{p}^{(n)} \subseteq \mathfrak{q}_n$ so that by Theorem 6, $\mathfrak{q}_n = \mathfrak{p}^{(m)}$ where $1 \leq m \leq n$, and therefore, $W(\mathfrak{q}_n) = W(\mathfrak{p}^{(m)}) = W(\mathfrak{p}^m) = mW(\mathfrak{p})$. From this, it is easy to see that the set of all $W(a)$ where $a \in \mathfrak{D}$ and $W(a) \neq \infty$, is just the ideal of the ring of ordinary integers generated by $W(\mathfrak{p})$, and since this ideal must be the whole ring of ordinary integers and $W(\mathfrak{p}) > 0$, then $W(\mathfrak{p}) = 1$ so that there exists an element $a \in \mathfrak{o}$ such that $W(a) = 1$. Consequently, for each positive integer n, $W(a^n) = n$ so that the $\mathfrak{q}_n$ are all distinct and therefore, $\mathfrak{q}_n = \mathfrak{p}^{(n)}$. By Theorem 3, this means that for all $a \in \mathfrak{o}$, $W(a) = V(a)$, and since $\mathfrak{o}$ is an order of $\mathfrak{D}$, one can easily show that for all $a \in \mathfrak{D}$, $W(a) = V(a)$, i.e. $W = V$.

THEOREM 9. *If $\{W_j\}_{j \in J}$ is a set of special valuations of $\mathfrak{D}$ such that for each regular element $a \in \mathfrak{D}$, $W_j(a) = 0$ for all $j \in J$ except a finite number, and if $\mathfrak{o}$ is the set of all $a \in \mathfrak{D}$ such that $W_j(a) \geq 0$ for all $j \in J$, then $\mathfrak{o}$ is a subring of $\mathfrak{D}$ containing the unity element of $\mathfrak{D}$. If $\mathfrak{o}$ is an order of $\mathfrak{D}$, then $\mathfrak{o}$ is a factorization ring and if $\{\mathfrak{p}_i\}_{i \in I}$ is the set of relevant prime ideals of $\mathfrak{o}$, and if for each $i \in I$, V_i is the valuation of $\mathfrak{D}$ determined by $\mathfrak{p}_i$, then $\{V_i\}_{i \in I}$ is a subset of $\{W_j\}_{j \in J}$.*

Proof.[2] Since $\mathfrak{o}$ is the intersection of the orders of $\mathfrak{D}$ determined by the W_j, $\mathfrak{o}$ is surely a subring of $\mathfrak{D}$ containing the unity element of $\mathfrak{D}$.

Now, let us suppose that $\mathfrak{o}$ is an order of $\mathfrak{D}$. If $a \in \mathfrak{D}$ and if b is a regular element of $\mathfrak{o}$ such that for all natural numbers n, $ba^n \in \mathfrak{o}$, this means that for all $j \in J$,

$$W_j(b) + nW_j(a) = W_j(ba^n) \geq 0$$

or $W_j(b) \geq n(-W_j(a))$, and since $\infty \neq W_j(b) \geq 0$, $-W_j(a) \leq 0$ and $W_j(a) \geq 0$ so that $a \in \mathfrak{o}$. Thus, $\mathfrak{o}$ is fully integrally closed in $\mathfrak{D}$.

If $\mathfrak{a}, \mathfrak{b} \in \mathfrak{L}(\mathfrak{o})$ and if $W_j(\mathfrak{a}) \leq W_j(\mathfrak{b})$ for all $j \in J$, then by Lemma 3,

$$W_j(\mathfrak{a}^{-1}\mathfrak{b}) = W_j(\mathfrak{a}^{-1}) + W_j(\mathfrak{b}) \geq W_j(\mathfrak{a}^{-1}) + W_j(\mathfrak{a}) = W_j(\mathfrak{a}\mathfrak{a}^{-1}) \geq 0$$

[2] This proof is a slightly modified version of a proof given by my student M. Aubert Daigneault for the case where $\mathfrak{o}$ is an integral domain, in his Master's thesis (Université de Montréal) entitled "Les anneaux de Dedekind."

so that $\mathfrak{a}^{-1}\mathfrak{b} \subseteq \mathfrak{o}$ and $\mathfrak{a} \leqslant \mathfrak{b}$. Therefore, $W_j(\mathfrak{a}) = W_j(\mathfrak{b})$ for all $j \in J$ implies that $\mathfrak{a} \sim \mathfrak{b}$.

Furthermore, if $\mathfrak{a} \in \mathfrak{F}(\mathfrak{o})$, then $W_j(\mathfrak{a}) = 0$ for all $j \in J$ except a finite number, for there exist regular elements $a, b \in \mathfrak{o}$ such that $a \in \mathfrak{a}$ and $b\mathfrak{a} \subseteq \mathfrak{o}$ and then, for all $j \in J$, $W_j(a) \geqslant W_j(\mathfrak{a}) \geqslant W_j(b^{-1})$ and $W_j(a) = W_j(b^{-1}) = 0$ for all $j \in J$ except a finite number.

Now, if $\overline{\mathfrak{a}_1} > \overline{\mathfrak{a}_2} > \overline{\mathfrak{a}_3} > \dots$ is a chain of elements of $\overline{\mathfrak{F}(\mathfrak{o})}$ where $\overline{\mathfrak{a}_k} \geqslant \overline{\mathfrak{o}}$ for all indices k, then $\mathfrak{a}_1^* \subset \mathfrak{a}_2^* \subset \mathfrak{a}_3^* \subset \dots$ and therefore, for all $j \in J$,

$$W_j(\mathfrak{a}_1^*) \geqslant W_j(\mathfrak{a}_2^*) \geqslant W_j(\mathfrak{a}_3^*) \geqslant \dots \geqslant 0.$$

Then, for each index k, $W_j(\mathfrak{a}_k^*) \neq W_j(\mathfrak{a}_{k+1}^*)$ for at least one $j \in J$ for otherwise, by what we have seen above, $\mathfrak{a}_k^* \sim \mathfrak{a}_{k+1}^*$, contradicting our hypothesis. Then, since $W_j(\mathfrak{a}_1^*) = 0$ for all $j \in J$ except a finite number, it is evident that the chain considered must be finite. Thus, $\mathfrak{o}$ is a factorization ring.

By Lemma 1, for each $j \in J$,

$$\mathfrak{P}_j = \{a \in \mathfrak{o} | W_j(a) > 0\}$$

is a regular prime ideal of $\mathfrak{o}$. If $\mathfrak{p}_i$ is a relevant prime ideal of $\mathfrak{o}$, $\mathfrak{p}_i$ contains a regular element a which is not a unit of $\mathfrak{o}$ so that $W_j(a) > 0$ for at least one $j \in J$. Let J_a denote the finite set of those indices $j \in J$ for which $W_j(a) > 0$, and for each $j \in J$, let $\mathfrak{q}_j$ denote the set of all $b \in \mathfrak{o}$ for which $W_j(b) \geqslant W_j(a)$. Then, $\mathfrak{q}_j \neq \mathfrak{o}$ if and only if $j \in J_a$ and by Lemma 1, for each $j \in J_a$, $\mathfrak{q}_j$ is $\mathfrak{P}_j$-primary. It is evident that

$$a \in \bigcap_{j \in J_a} \mathfrak{q}_j = \bigcap_{j \in J} \mathfrak{q}_j .$$

If b is in the intersection of all the $\mathfrak{q}_j$, for all $j \in J$, $W_j(b) \geqslant W_j(a)$ so that $W_j(ba^{-1}) = W_j(b) - W_j(a) \geqslant 0$ and therefore, $ba^{-1} \in \mathfrak{o}$ and $b \in (a)$. Therefore,

$$(a) = \bigcap_{j \in J_a} \mathfrak{q}_j .$$

Since $\mathfrak{p}_i \supseteq (a)$, $\mathfrak{p}_i$ contains one of the $\mathfrak{P}_j$ with $j \in J_a$, say $\mathfrak{p}_i \supseteq \mathfrak{P}_{j_1}$. Then, since $\mathfrak{P}_{j_1}$ is a regular prime ideal of $\mathfrak{o}$ and $\mathfrak{p}_1$ is a minimal regular prime ideal of $\mathfrak{o}$, $\mathfrak{p}_i = \mathfrak{P}_{j_1}$ and by Lemma 4, $V_i = W_{j_1}$.

THEOREM 10. *If V is a special valuation of $\mathfrak{D}$ and if $\mathfrak{o}$ is the order of $\mathfrak{D}$ determined by V, then $V_\mathfrak{o}$ induces an isomorphism of $\overline{\mathfrak{L}(\mathfrak{o})}$ onto G'', this isomorphism mapping $\overline{\mathfrak{F}(\mathfrak{o})}$ onto G.*

Proof. By Lemma 3, $V_\mathfrak{o}$ is a homomorphism of $\mathfrak{L}(\mathfrak{o})$ onto G''. If $\mathfrak{a}, \mathfrak{b} \in \mathfrak{L}(\mathfrak{o})$,

$$\begin{aligned}
\mathfrak{a} \leqslant \mathfrak{b} &\leftrightarrow \mathfrak{a}^{-1} \subseteq \mathfrak{b}^{-1} \\
&\leftrightarrow (a \in \mathfrak{D})(a\mathfrak{a} \subseteq \mathfrak{o} \rightarrow a\mathfrak{b} \subseteq \mathfrak{o}) \\
&\leftrightarrow (a \in \mathfrak{D})(V(a\mathfrak{a}) \geqslant 0 \rightarrow V(a\mathfrak{b}) \geqslant 0) \\
&\leftrightarrow (a \in \mathfrak{D})(V(a) + V(\mathfrak{a}) \geqslant 0 \rightarrow V(a) + V(\mathfrak{b}) \geqslant 0) \\
&\leftrightarrow V(\mathfrak{a}) \leqslant V(\mathfrak{b}).
\end{aligned}$$

Therefore, also $\mathfrak{a} \sim \mathfrak{b}$ if and only if $V(\mathfrak{a}) = V(\mathfrak{b})$ and V_* induces an isomorphism of $\overline{\mathfrak{L}(\mathfrak{o})}$ onto G''.

By Theorem 9, $\mathfrak{o}$ is a factorization ring with a single relevant prime ideal $\mathfrak{p}$ and V is the valuation of $\mathfrak{D}$ determined by $\mathfrak{p}$. It is evident that for each non-negative integer n, $V(\mathfrak{p}^n) = n$. Since $\mathfrak{p}^n(\mathfrak{p}^n)^{-1} \subseteq \mathfrak{o}$, $V(\mathfrak{p}^n(\mathfrak{p}^n)^{-1}) > 0$. If we suppose that $V(\mathfrak{p}^n(\mathfrak{p}^n)^{-1}) > 0$, then $\mathfrak{p}^n(\mathfrak{p}^n)^{-1} \subseteq \mathfrak{p}$ so that $\mathfrak{p}^n(\mathfrak{p}^n)^{-1} > \mathfrak{p}$ and $\mathfrak{p}^n > \mathfrak{p}^{n+1}$, a contradiction. Therefore,

$$V(\mathfrak{p}^n) + V((\mathfrak{p}^n)^{-1}) = V(\mathfrak{p}^n(\mathfrak{p}^n)^{-1}) = 0$$

so that $V((\mathfrak{p}^n)^{-1}) = -n$. Therefore, V_* maps $\mathfrak{F}(\mathfrak{o})$ onto G.

In Lemmas 5, 6, 7 and 8, $\mathfrak{o}$ is to be considered as an order of $\mathfrak{D}$ with the property that for each regular element $a \in \mathfrak{o}$ that is not a unit of $\mathfrak{o}$, (a) is the intersection of a finite number of formal powers of minimal regular prime ideals of $\mathfrak{o}$. In all of these lemmas, we may set aside the trivial case where $\mathfrak{o} = \mathfrak{D}$ so that the minimal regular prime ideals of $\mathfrak{o}$ are all proper.

LEMMA 5. *If $\mathfrak{p}$ is a minimal regular prime ideal of $\mathfrak{o}$, then for each positive integer n, $\mathfrak{p}^{(n)} \sim \mathfrak{p}^n$.*

Proof. Since $\mathfrak{p}^n \subseteq \mathfrak{p}^{(n)}$, $(\mathfrak{p}^n)^{-1} \supseteq (\mathfrak{p}^{(n)})^{-1}$. If $a \in (\mathfrak{p}^n)^{-1}$, $a = b/c$ where $b, c \in \mathfrak{o}$ and c is regular. If c is a unit of $\mathfrak{o}$, then $a \in \mathfrak{o} \subseteq (\mathfrak{p}^{(n)})^{-1}$. If c is not a unit of $\mathfrak{o}$,

$$(c) = \mathfrak{p}_1^{(n_1)} \cap \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)},$$

where $\mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_r$ are distinct minimal regular prime ideals of $\mathfrak{o}$. Since $bc^{-1}\mathfrak{p}^n \subseteq \mathfrak{o}$, $b\mathfrak{p}^n \subseteq (c)$. If $\mathfrak{p}$ is different from each $\mathfrak{p}_i$, then

$$b\mathfrak{p}^n \subseteq \mathfrak{p}_i^{(n_i)} \;\&\; \mathfrak{p}^n \not\subseteq \mathfrak{p}_i \rightarrow b \in \mathfrak{p}_i^{(n_i)} \qquad (i \leqslant r)$$

and therefore, $b \in (c)$, $a = b/c \in \mathfrak{o} \subseteq (\mathfrak{p}^{(n)})^{-1}$. If $\mathfrak{p}$ is equal to one of the $\mathfrak{p}_i$, say $\mathfrak{p} = \mathfrak{p}_1$, then, by the same argument as above,

$$b \in \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)}.$$

Then, if $d \in \mathfrak{p}^{(n)}$, there exists $g \in \mathfrak{o}$ such that $g \notin \mathfrak{p}$ and $gd \in \mathfrak{p}^n$ so that

$$g(bd) = b(gd) \in (c) \subseteq \mathfrak{p}^{(n_1)};$$

also

$$g \notin \mathfrak{p} \rightarrow bd \in \mathfrak{p}^{(n_1)} \rightarrow b\mathfrak{p}^{(n)} \subseteq \mathfrak{p}^{(n_1)}$$

and therefore,

$$b\mathfrak{p}^{(n)} \subseteq \mathfrak{p}_1^{(n_1)} \cap \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)} = (c).$$

Therefore, $a\mathfrak{p}^{(n)} = bc^{-1}\mathfrak{p}^{(n)} \subseteq \mathfrak{o}$ so that $a \in (\mathfrak{p}^{(n)})^{-1}$ and $(\mathfrak{p}^n)^{-1} = (\mathfrak{p}^{(n)})^{-1}$.

LEMMA 6. *If $\mathfrak{p}_1$ and $\mathfrak{p}_2$ are two distinct minimal regular prime ideals of $\mathfrak{o}$ and if k_1 and k_2 are two positive integers, then*

$$\mathfrak{p}_1^{(k_1)} + \mathfrak{p}_2^{(k_2)} \sim \mathfrak{o}.$$

Proof.

$$\mathfrak{p}_1^{(k_1)} + \mathfrak{p}_2^{(k_2)} \subseteq \mathfrak{o} \rightarrow (\mathfrak{p}_1^{(k_1)} + \mathfrak{p}_2^{(k_2)})^{-1} \supseteq \mathfrak{o}.$$

If

$$a \in (\mathfrak{p}_1^{(k_1)} + \mathfrak{p}_2^{(k_2)})^{-1},$$

$a = b/c$ where $b, c \in \mathfrak{o}$ and c is regular, and

$$a(\mathfrak{p}_1^{(k_1)} + \mathfrak{p}_2^{(k_2)}) \subseteq \mathfrak{o} \rightarrow a\mathfrak{p}_r^{(k_r)} \subseteq \mathfrak{o} \qquad (r = 1,2)$$
$$\rightarrow b\mathfrak{p}_r^{(k_r)} \subseteq (c).$$

Of course, if c is a unit of $\mathfrak{o}$, then $a \in \mathfrak{o}$ and there is nothing more to prove so that we may suppose that c is not a unit of $\mathfrak{o}$ and therefore,

$$(c) = \mathfrak{p}_{i_1}^{(n_1)} \cap \mathfrak{p}_{i_2}^{(n_2)} \cap \ldots \cap \mathfrak{p}_{i_r}^{(n_r)}$$

where

$$\mathfrak{p}_{i_1}, \mathfrak{p}_{i_2}, \ldots, \mathfrak{p}_{i_r},$$

are distinct minimal regular prime ideals of $\mathfrak{o}$. If $\mathfrak{p}_1$ is different from each $\mathfrak{p}_{i_j}$, then

$$b\mathfrak{p}_1^{(k_1)} \subseteq \mathfrak{p}_{i_j}^{(n_j)} \;\&\; \mathfrak{p}_1^{(k_1)} \not\subseteq \mathfrak{p}_{i_j} \rightarrow b \in \mathfrak{p}_{i_j}^{(n_j)} \qquad (j \leqslant r),$$

and therefore, $b \in (c)$ and $a = b/c \in \mathfrak{o}$. Similarly, if $\mathfrak{p}_2$ is different from each $\mathfrak{p}_{i_j}$, one may show that $a \in \mathfrak{o}$.

Now, if

$$\mathfrak{p}_1 = \mathfrak{p}_{i_1}, \mathfrak{p}_2 = \mathfrak{p}_{i_2},$$

then

$$b\mathfrak{p}_1^{(k_1)} \subseteq \mathfrak{p}_{i_j}^{(n_j)} \;\&\; \mathfrak{p}_1^{(k_1)} \not\subseteq \mathfrak{p}_{i_j} \rightarrow b \in \mathfrak{p}_{i_j}^{(n_j)} \qquad (1 < j \leqslant r)$$

and

$$b\mathfrak{p}_2^{(k_2)} \subseteq \mathfrak{p}_{i_1}^{(n_1)} \;\&\; \mathfrak{p}_2^{(k_2)} \not\subseteq \mathfrak{p}_{i_1} \rightarrow b \in \mathfrak{p}_{i_1}^{(n_1)},$$

so that $b \in (c)$ and $a = b/c \in \mathfrak{o}$. Therefore

$$(\mathfrak{p}_1^{(k_1)} + \mathfrak{p}_2^{(k_2)}) = \mathfrak{o}.$$

LEMMA 7. *If $\mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_r$ are distinct minimal regular prime ideals of $\mathfrak{o}$ and if $n_1, n_2, \ldots, n_r$ are positive integers, then*

$$\mathfrak{p}_1^{(n_1)} \cap \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)} \sim \mathfrak{p}_1^{n_1}\mathfrak{p}_2^{n_2} \ldots \mathfrak{p}_r^{n_r}$$

Proof. By induction from Lemmas 5 and 6, using Propositions 3 and 4.

LEMMA 8. *If $\mathfrak{p}$ is a minimal regular prime ideal of $\mathfrak{o}$, then $\mathfrak{p} > \mathfrak{o}$, $\bar{\mathfrak{p}}$ is invertible and for each positive integer n, $(\mathfrak{p}^n)^* = \mathfrak{p}^{(n)}$.*

Proof. Let a be a regular element of $\mathfrak{o}$ contained in $\mathfrak{p}$. Then,

$$(a) = \mathfrak{p}_1^{(n_1)} \cap \mathfrak{p}_2^{(n_2)} \cap \ldots \cap \mathfrak{p}_r^{(n_r)}$$

where $\mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_r$ are distinct minimal regular prime ideals of $\mathfrak{o}$, and $\mathfrak{p}$ must coincide with one of the $\mathfrak{p}_i$, say $\mathfrak{p} = \mathfrak{p}_1$. By Lemma 7,

$$(a) \sim \mathfrak{p}_1^{n_1}\mathfrak{p}_2^{n_2} \ldots \mathfrak{p}_r^{n_r}$$

so that

$$\mathfrak{p}_1(\mathfrak{p}_1^{n_1-1}\mathfrak{p}_2^{n_2} \ldots \mathfrak{p}_r^{n_r}a^{-1}) \sim \mathfrak{o}$$

i.e., $\overline{\mathfrak{p}}$ is invertible.

If $\mathfrak{p} \sim \mathfrak{o}$, then

$$\mathfrak{p}_1^{n_1} \sim \mathfrak{o} \text{ so that } (a) \sim \mathfrak{p}_2^{n_2} \ldots \mathfrak{p}_r^{n_r}$$

and since $(a) = (a)^*$,

$$\mathfrak{p} \supseteq (a) \supseteq \mathfrak{p}_2^{n_2} \ldots \mathfrak{p}_r^{n_r}$$

so that $\mathfrak{p} = \mathfrak{p}_1$ contains some $\mathfrak{p}_i$ with $i \neq 1$, contradicting the hypothesis that the $\mathfrak{p}_i$ are all distinct. Therefore, $\mathfrak{p} > \mathfrak{o}$.

Since $\overline{\mathfrak{p}}$ is invertible, for each natural number n, $\overline{\mathfrak{p}^n}$ is invertible and by Proposition 6, $\mathfrak{p}^n:(\mathfrak{p}^n)^* \sim \mathfrak{p}^n(\mathfrak{p}^n)^{-1} \sim \mathfrak{o}$ so that there exists $b \in \mathfrak{o}$ such that $b \notin \mathfrak{p}$ and $b(\mathfrak{p}^n)^* \subseteq \mathfrak{p}^n \subseteq \mathfrak{p}^{(n)}$ and therefore, $(\mathfrak{p}^n)^* \subseteq \mathfrak{p}^{(n)}$ and by Lemma 5, $(\mathfrak{p}^n)^* = \mathfrak{p}^{(n)}$.

THEOREM 11. *If $\mathfrak{o}$ is an order of $\mathfrak{O}$ with the property that for each regular element $a \in \mathfrak{o}$ that is not a unit of $\mathfrak{o}$, (a) is the intersection of a finite number of formal powers of minimal regular prime ideals of $\mathfrak{o}$, then $\mathfrak{o}$ is a factorization ring.*

Proof. Let $\{\mathfrak{p}_i\}_{i \in I}$ be the set of all minimal regular prime ideals of $\mathfrak{o}$. For each $i \in I$, because of the properties of $\mathfrak{p}_i$ given by Lemma 8, we may define a special valuation V_i of $\mathfrak{O}$ in exactly the same way as we defined the valuation of $\mathfrak{O}$ determined by a relevant prime ideal of a factorization ring having $\mathfrak{O}$ as full ring of quotients in the preceding section. Then, if $a \in \mathfrak{o}$, it is evident that for all $i \in I$, $V_i(a) \geqslant 0$. Conversely, if $a \in \mathfrak{O}$ and if $V_i(a) \geqslant 0$ for all $i \in I$, then $a = b/c$ where $b, c \in \mathfrak{o}$ and c is regular, and if c is not a unit of $\mathfrak{o}$, then

$$(c) = \mathfrak{p}_{i_1}^{(n_1)} \cap \mathfrak{p}_{i_2}^{(n_2)} \cap \ldots \cap \mathfrak{p}_{i_r}^{(n_r)}$$

where $i_1, i_2, \ldots, i_r$ are distinct elements of I, and one can easily verify that for $j \leqslant r$, $V_{i_j}(c) = n_j$, while for $i \in I$, $i \neq i_j$, $V_i(c) = 0$. Then,

$$V_{i_j}(a) = V_{i_j}(b) - V_{i_j}(c) \geqslant 0 \rightarrow V_{i_j}(b) \geqslant V_{i_j}(c) \rightarrow b \in \mathfrak{p}_{i_j}^{(n_j)} \quad (j \leqslant r)$$

and therefore, $b \in (c)$ and $a = b/c \in \mathfrak{o}$. Therefore, by Theorem 9, $\mathfrak{o}$ is a factorization ring.

6. Generalized Dedekind rings. We will say that a factorization ring $\mathfrak{o}$ is a generalized Dedekind ring, if for $\underline{\mathfrak{a}}, \underline{\mathfrak{b}} \in \mathfrak{F}(\mathfrak{o})$, $\mathfrak{a} \sim \mathfrak{b}$ implies that $\mathfrak{a} = \mathfrak{b}$. This means that $\mathfrak{F}(\mathfrak{o})$ is isomorphic to $\overline{\mathfrak{F}(\mathfrak{o})}$ so that every proper regular ideal

of $\mathfrak{o}$ may be expressed in a unique way as the product of a finite number of relevant prime ideals of $\mathfrak{o}$.

If $\mathfrak{o}$ is an order of $\mathfrak{D}$, then $\mathfrak{o}$ is a generalized Dedekind ring if and only if $\mathfrak{o}$ satisfies one of the following sets of conditions:

1. $\mathfrak{F}(\mathfrak{o})$ is a group.
2. For any two regular ideals $\mathfrak{a}$ and $\mathfrak{b}$ of $\mathfrak{o}$, $\mathfrak{a} \subseteq \mathfrak{b}$ implies that there exists an ideal $\mathfrak{c}$ of $\mathfrak{o}$ (it must be regular) such that $\mathfrak{a} = \mathfrak{b}\mathfrak{c}$.
3. (i) $\mathfrak{o}$ is a factorization ring,
   (ii) the relevant prime ideals of $\mathfrak{o}$ are the only proper regular prime ideals of $\mathfrak{o}$.
4. (i) $\mathfrak{o}$ is fully integrally closed in $\mathfrak{D}$,
   (ii) if $\mathfrak{a}$ is any regular ideal of $\mathfrak{o}$, then the descending chain condition is valid for the ideals of $\mathfrak{o}/\mathfrak{a}$.

We will not prove the equivalence of these sets of conditions, the proofs being entirely similar to those given in the case of integral domains.

For simplicity, we will use the term "Dedekind ring" instead of "generalized Dedekind ring" and "Dedekind domain" instead of "ordinary Dedekind ring."

THEOREM 12. *If $\mathfrak{o}$ is a Dedekind ring with $\mathfrak{D}$ as full ring of quotients and if V is a special valuation of $\mathfrak{D}$ with the property that $V(a) \geqslant 0$ for all $a \in \mathfrak{o}$, then V is determined by a relevant prime ideal of $\mathfrak{o}$.*

Proof. By Lemma 1, $\mathfrak{p} = \{a \in \mathfrak{o} | V(a) > 0\}$ is a proper regular prime ideal of $\mathfrak{o}$, and since $\mathfrak{o}$ is a Dedekind ring, $\mathfrak{p}$ is a relevant prime ideal of $\mathfrak{o}$ so that by Lemma 4, V is determined by $\mathfrak{p}$.

THEOREM 13. *If V is a special valuation of $\mathfrak{D}$, if $\mathfrak{o}$ is the order of $\mathfrak{D}$ determined by V and if $\mathfrak{D}$ satisfies either one of the following two conditions:*

1. *there is only a finite number of maximal prime ideals of* (0) *in* $\mathfrak{D}$,
2. $\mathfrak{D}$ *is "einartig"* (**4**, p. 22),

then $\mathfrak{o}$ is a Dedekind ring.

Proof. Let $\mathfrak{p}$ denote the unique relevant prime ideal of $\mathfrak{o}$.

1. Let us assume that there is only a finite number of maximal prime ideals $\mathfrak{P}_1, \mathfrak{P}_2, \ldots, \mathfrak{P}_n$ of (0) in $\mathfrak{D}$ (this condition is satisfied when the ascending chain condition holds for the ideals of $\mathfrak{D}$ and a fortiori, when it holds for the ideals of $\mathfrak{o}$). Let $\mathfrak{a}$ be a proper ideal of $\mathfrak{o}$. If $a \in \mathfrak{a}$ and $a \notin \mathfrak{p}$, then $V(a) = 0$ and a is not a regular element of $\mathfrak{o}$, for if a were regular, $V(a^{-1}) = -V(a) = 0$ so that $a^{-1} \in \mathfrak{o}$ and $\mathfrak{a}$ would not be a proper ideal of $\mathfrak{o}$. Then,

$$a \in (\mathfrak{P}_1 \cup \mathfrak{P}_2 \cup \ldots \cup \mathfrak{P}_n) \cap \mathfrak{o} = (\mathfrak{P}_1 \cap \mathfrak{o}) \cup (\mathfrak{P}_2 \cap \mathfrak{o}) \cup \ldots \cup (\mathfrak{P}_n \cap \mathfrak{o})$$

and therefore,

$$\mathfrak{a} \subseteq \mathfrak{p} \cup (\mathfrak{P}_1 \cap \mathfrak{o}) \cup (\mathfrak{P}_2 \cap \mathfrak{o}) \cup \ldots \cup (\mathfrak{P}_n \cap \mathfrak{o})$$

Since the ideals $\mathfrak{p}, \mathfrak{P}_1 \cap \mathfrak{o}, \mathfrak{P}_2 \cap \mathfrak{o}, \ldots, \mathfrak{P}_n \cap \mathfrak{o}$ are prime ideals of $\mathfrak{o}$, $\mathfrak{a}$

must be contained in one of them. But if $\mathfrak{a} \subseteq (\mathfrak{P}_i \cap \mathfrak{o})$, $\mathfrak{a}$ is not regular. Therefore, $\mathfrak{p}$ is the only proper regular prime ideal of $\mathfrak{o}$.

2. Let us assume that $\mathfrak{O}$ is "einartig" (this condition is certainly satisfied when the descending chain condition is valid for the ideals of $\mathfrak{O}$). By Lemma 1, the associate $\mathfrak{p}'$ of $\mathfrak{p}$ is a proper prime ideal of $\mathfrak{O}$ so that it must be a maximal proper ideal of $\mathfrak{O}$ and $\mathfrak{O}/\mathfrak{p}'$ is a field. Since $\mathfrak{p}'$ is the kernel of V, we may speak of the projection V' of V by the natural homomorphism of $\mathfrak{O}$ onto $\mathfrak{O}/\mathfrak{p}'$. Then, $\mathfrak{o}/\mathfrak{p}'$ is the order of $\mathfrak{O}/\mathfrak{p}'$ determined by V' (rank 1, discrete valuation of the field $\mathfrak{O}/\mathfrak{p}'$) so that $\mathfrak{p}/\mathfrak{p}'$ is the only non-trivial prime ideal of $\mathfrak{o}/\mathfrak{p}'$ and $\mathfrak{p}$ is the only proper regular prime ideal of $\mathfrak{o}$.

For the remainder of this section, we will assume that $\mathfrak{o}$ is a Dedekind ring and that the ascending chain condition is valid for the ideals of $\mathfrak{o}$. We will denote the set of all regular elements of $\mathfrak{o}$ by S, $\{\mathfrak{p}_i\}_{i \in I}$ will be the set of relevant prime ideals of $\mathfrak{o}$ and for each $i \in I$, V_i will denote the valuation of $\mathfrak{O} = \mathfrak{o}_S$ determined by $\mathfrak{p}_i$. If $\mathfrak{a}$ is an ideal of $\mathfrak{o}$, $I(\mathfrak{a})$ will denote the set of all $i \in I$ for which $0 < V_i(\mathfrak{a}) < \infty$ and $I'(\mathfrak{a})$ will denote the set of all $i \in I$ for which $V_i(\mathfrak{a}) = \infty$. If $i \in I$, since

$$\mathfrak{p}_i' = \bigcap_{n=1}^{\infty} \mathfrak{p}_i^{(n)} = \bigcap_{n=1}^{\infty} \mathfrak{p}_i^n$$

we will set $\mathfrak{p}_i' = \mathfrak{p}_i^{(\infty)} = \mathfrak{p}_i^{\infty}$.

THEOREM 14. *If $\mathfrak{a}$ is an ideal of $\mathfrak{o}$, then $I(\mathfrak{a})$ is finite, for each $i \in I(\mathfrak{a})$, $\mathfrak{a}_S$ and $\mathfrak{p}_i$ are without proper common divisor and*

$$\mathfrak{a} = \mathfrak{a}_S \cdot \prod_{i \in I(\mathfrak{a})} \mathfrak{p}_i^{m_i}$$

where $m_i = V_i(\mathfrak{a})$. We will call this representation of $\mathfrak{a}$ the standard decomposition of $\mathfrak{a}$.

Proof. Let $\mathfrak{a} = \mathfrak{q}_1 \cap \mathfrak{q}_2 \cap \ldots \cap \mathfrak{q}_n$ be a normal decomposition of $\mathfrak{a}$ into primary ideals. The radicals of some of these primary ideals may be relevant prime ideals of $\mathfrak{o}$. Let us say that the radicals of $\mathfrak{q}_1, \mathfrak{q}_2, \ldots, \mathfrak{q}_r$ $(0 \leqslant r \leqslant n)$ are the relevant prime ideals

$$\mathfrak{p}_{i_1}, \mathfrak{p}_{i_2}, \ldots, \mathfrak{p}_{i_r}$$

respectively $(i_j \in I)$, while the radicals of $\mathfrak{q}_{r+1}, \ldots, \mathfrak{q}_n$ are not relevant prime ideals of $\mathfrak{o}$. Since the relevant prime ideals of $\mathfrak{o}$ are the only proper regular prime ideals of $\mathfrak{o}$,

$$\mathfrak{a}_S = \mathfrak{q}_{r+1} \cap \ldots \cap \mathfrak{q}_n.$$

By Theorem 6, for each $k \leqslant r$, $\mathfrak{q}_k$ is a formal power, say

$$\mathfrak{q}_k = \mathfrak{p}_{i_k}^{(n_k)},$$

and since $\mathfrak{o}$ is a Dedekind ring,

$$\mathfrak{q}_1 \cap \mathfrak{q}_2 \cap \ldots \cap \mathfrak{q}_r = \mathfrak{p}_{i_1}^{(n_1)} \cap \mathfrak{p}_{i_2}^{(n_2)} \cap \ldots \cap \mathfrak{p}_{i_r}^{(n_r)} = \mathfrak{p}_{i_1}^{n_1} \mathfrak{p}_{i_2}^{n_2} \ldots \mathfrak{p}_{i_r}^{n_r}.$$

Now, if $h > r$, then the radical of $\mathfrak{q}_h$ is not contained in any of the ideals

$$\mathfrak{p}_{i_1}, \mathfrak{p}_{i_2}, \ldots, \mathfrak{p}_{i_r},$$

for if

$$\operatorname{rad} \mathfrak{q}_h \subseteq \mathfrak{p}_{i_k} \qquad (1 \leqslant k \leqslant r),$$

since $\operatorname{rad} \mathfrak{q}_h$ is a prime ideal of $\mathfrak{o}$, but not a relevant prime ideal of $\mathfrak{o}$, by Theorem 4,

$$\operatorname{rad} \mathfrak{q}_h \subseteq \mathfrak{p}'_{i_k}, \mathfrak{q}_k = \mathfrak{p}_{i_k}^{n_k} \supseteq \mathfrak{p}'_{i_k} \supseteq \mathfrak{q}_h,$$

contradicting the hypothesis that the given decomposition of $\mathfrak{a}$ into primary ideals is normal. From this, we deduce first of all, since the $\mathfrak{p}$'s are maximal proper ideals of $\mathfrak{o}$, that $\mathfrak{a}_S$ and each $\mathfrak{p}_{i_k}$ are without proper common divisors, so that

$$\mathfrak{a} = \mathfrak{p}_{i_1}^{n_1}\mathfrak{p}_{i_2}^{n_2} \ldots \mathfrak{p}_{i_r}^{n_r} \cap \mathfrak{a}_S = \mathfrak{p}_{i_1}^{n_1}\mathfrak{p}_{i_2}^{n_2} \ldots \mathfrak{p}_{i_r}^{n_r}\mathfrak{a}_S.$$

Secondly, for each $k \leqslant r$, $\mathfrak{p}_{i_k}$ is a minimal prime ideal of $\mathfrak{a}$ and

$$\mathfrak{a}_{\mathfrak{p}_{i_k}} = \mathfrak{p}_{i_k}^{n_k} \to \mathfrak{a} \subseteq \mathfrak{p}_{i_k}^{n_k} \text{ and } \mathfrak{a} \not\subseteq \mathfrak{p}_{i_k}^{n_k+1}$$
$$\to n_k = V_{i_k}(\mathfrak{a}) = m_k \text{ and } \{i_1, i_2, \ldots, i_r\} \subseteq I(\mathfrak{a}).$$

Now, let us assume that $i \in I$ and that $i \notin \{i_1, i_2, \ldots, i_r\}$. First of all, if $\mathfrak{a} \not\subseteq \mathfrak{p}_i$, then $V_i(\mathfrak{a}) = 0$ and $i \notin I(\mathfrak{a})$. Secondly, if $\mathfrak{a} \subseteq \mathfrak{p}_i$, since $\mathfrak{p}_i$ does not belong to $\mathfrak{a}$, all prime ideals belonging to $\mathfrak{a}$ and contained in $\mathfrak{p}_i$ must be contained in $\mathfrak{p}_i'$ (Theorem 4). Since $\mathfrak{p}_i'$ is a minimal prime ideal of $\mathfrak{o}$, $\mathfrak{p}_i'$ is the only prime ideal belonging to $\mathfrak{a}$ and contained in $\mathfrak{p}_i$ so that $V_i(\mathfrak{a}) = \infty$ and $i \notin I(\mathfrak{a})$.

Theorem 14 implies that an ideal $\mathfrak{a}$ of $\mathfrak{o}$ is completely determined by its isolated component $\mathfrak{a}_S$ and by its values $V_i(\mathfrak{a})$, for all $i \in I(\mathfrak{a})$.

Let us notice that if $\mathfrak{a}$ and $\mathfrak{b}$ are ideals of $\mathfrak{o}$, then by Lemma 3,

$$V_i(\mathfrak{ab}) = V_i(\mathfrak{a}) + V_i(\mathfrak{b}),$$
$$V_i(\mathfrak{a} + \mathfrak{b}) = \min\{V_i(\mathfrak{a}), V_i(\mathfrak{b})\},$$

and by one of the remarks made after Theorem 7, if $V_i(\mathfrak{a}) = m_i$ and $V_i(\mathfrak{b}) = n_i$, then

$$(\mathfrak{a} \cap \mathfrak{b})_{\mathfrak{p}_i} = \mathfrak{a}_{\mathfrak{p}_i} \cap \mathfrak{b}_{\mathfrak{p}_i} = \mathfrak{p}_i^{m_i} \cap \mathfrak{p}_i^{n_i} = \mathfrak{p}_i^{\max\{m_i, n_i\}},$$

so that $V_i(\mathfrak{a} \cap \mathfrak{b}) = \max\{V_i(\mathfrak{a}), V_i(\mathfrak{b})\}$. These rules are useful for finding $I(\mathfrak{ab})$, $I(\mathfrak{a} + \mathfrak{b})$ and $I(\mathfrak{a} \cap \mathfrak{b})$ when the sets $\{V_i(\mathfrak{a})\}_{i \in I}$ and $\{V_i(\mathfrak{b})\}_{i \in I}$ are given.

The standard decomposition of an ideal $\mathfrak{a}$ of $\mathfrak{o}$ is not the only representation of $\mathfrak{a}$ as the product of $\mathfrak{a}_S$ with a finite number of positive powers of relevant prime ideals of $\mathfrak{o}$. For example, if $\mathfrak{p}$ is a relevant prime ideal of $\mathfrak{o}$ and if $\mathfrak{p}'$ is the associate of $\mathfrak{p}$, then $\mathfrak{p}' = (\mathfrak{p}')_S$ is the standard decomposition of $\mathfrak{p}'$. But since

$$\mathfrak{p}' = \bigcap_{n=1}^{\infty} \mathfrak{p}^n$$

by Krull's Intersection Theorem, $\mathfrak{p}' = (0)_{S'}$, where S' is the set of all elements of $\mathfrak{o}$ that are of the form $1 - a$, where $a \in \mathfrak{p}$. Then, if $b \in \mathfrak{p}'$, there exists $a \in \mathfrak{p}$ such that $(1 - a)b = 0$ so that $b = ab \in \mathfrak{p}\mathfrak{p}'$ and $\mathfrak{p}' = \mathfrak{p}\mathfrak{p}' = \mathfrak{p}(\mathfrak{p}')_S$.

THEOREM 15. *If $\mathfrak{a}$ is an ideal of $\mathfrak{o}$ and if*

$$\mathfrak{a} = \mathfrak{a}_S \prod_{i \in I(\mathfrak{a})} \mathfrak{p}_i^{m_i}$$

is the standard decomposition of $\mathfrak{a}$, then any representation of $\mathfrak{a}$ as the product of $\mathfrak{a}_S$ with a finite number of positive powers of relevant prime ideals of $\mathfrak{o}$ is of the form

$$\mathfrak{a} = \mathfrak{a}_S \prod_{i \in I(\mathfrak{a})} \mathfrak{p}_i^{m_i} \prod_{i \in J} \mathfrak{p}_i^{n_i},$$

where J is a finite subset of $I'(\mathfrak{a})$.

Proof. Let us suppose that

$$\mathfrak{a}_S \prod_{i \in I(\mathfrak{a})} \mathfrak{p}_i^{m_i} = \mathfrak{a}_S \prod_{i \in K} \mathfrak{p}_i^{n_i}, \tag{1}$$

where K is a finite subset of I and $n_i > 0$ for all $i \in K$. If $j \in I(\mathfrak{a})$, by Theorem 14, $\mathfrak{p}_j$ and $\mathfrak{a}_S$ are without proper common divisor so that $\mathfrak{a}_S \not\subseteq \mathfrak{p}_j$, and since

$$\mathfrak{p}_j \supseteq \mathfrak{a}_S \prod_{i \in K} \mathfrak{p}_i^{n_i},$$

$\mathfrak{p}_j$ must contain, and therefore be equal to, some $\mathfrak{p}_i$ for $i \in K$. Since $\mathfrak{p}_j$ is invertible, we may cancel $\mathfrak{p}_j$ from both sides of (1). It is then evident that we may repeat this argument until all the relevant prime ideals appearing on the left-hand side of (1) have been cancelled so that $I(\mathfrak{a})$ is a subset of K, for each $i \in I(\mathfrak{a})$, $m_i \leqslant n_i$ and if we set $J = K - I(\mathfrak{a})$,

$$\mathfrak{a}_S = \mathfrak{a}_S \prod_{i \in I(\mathfrak{a})} \mathfrak{p}_i^{n_i - m_i} \prod_{i \in J} \mathfrak{p}_i^{n_i}.$$

From this equation, since for each $i \in I(\mathfrak{a})$, $\mathfrak{p}_i \not\supseteq \mathfrak{a}_S$, it is evident that $m_i = n_i$ so that

$$\mathfrak{a}_S = \mathfrak{a}_S \prod_{i \in J} \mathfrak{p}_i^{n_i}.$$

But if $i \in J$, $\mathfrak{p}_i \supseteq \mathfrak{a}$ and therefore, $V_i(\mathfrak{a}) > 0$. But then, since $i \notin I(\mathfrak{a})$, $V_i(\mathfrak{a}) = \infty$ and $i \in I'(\mathfrak{a})$.

7. A special case. In this section, we will assume that the descending chain condition is valid for the ideals of $\mathfrak{O}$ and our object will be to determine all orders of $\mathfrak{O}$ that are Noetherian Dedekind rings.

If $\mathfrak{N}_1, \mathfrak{N}_2, \ldots, \mathfrak{N}_p$ are the proper prime ideals of $\mathfrak{O}$, then

$$(0) = \bigcap_{h=1}^{p} \mathfrak{O}_h \qquad \text{(direct intersection)},$$

where each $\mathfrak{O}_h$ is $\mathfrak{N}_h$-primary, and

$$\mathfrak{O} = \sum_{k=1}^{p} \mathfrak{O}_k \qquad \text{(direct sum)},$$

where

$$\mathfrak{O}_k = \bigcap_{\substack{h=1 \\ h \neq k}}^{p} \mathfrak{O}_h \quad \text{and} \quad \mathfrak{O}_h = \sum_{\substack{k=1 \\ k \neq h}}^{p} \mathfrak{O}_k .$$

Also, $\mathfrak{O}_k \cong \mathfrak{O}/\mathfrak{O}_k$ so that $\mathfrak{O}_k$ is completely primary, i.e., every non-regular element of $\mathfrak{O}_k$ is nilpotent.

If for each $k = 1,2, \ldots, p$, $\mathfrak{o}_k$ is an order of $\mathfrak{O}_k$, then it is easy to show that

$$\mathfrak{o} = \sum_{k=1}^{p} \mathfrak{o}_k$$

is an order of $\mathfrak{O}$ and that $\mathfrak{o}$ is fully integrally closed in $\mathfrak{O}$ if and only if each $\mathfrak{o}_k$ is fully integrally closed in $\mathfrak{O}_k$. If $\mathfrak{o}$ is any order of $\mathfrak{O}$ and if Φ_k is the projection function of $\mathfrak{O}$ onto $\mathfrak{O}_k$, then $\Phi_k(\mathfrak{o})$ is an order of $\mathfrak{O}_k$ and if

$$\mathfrak{o} = \sum_{k=1}^{p} \Phi_k(\mathfrak{o})$$

then we will say that $\mathfrak{o}$ is decomposable.

LEMMA 9. *If $\mathfrak{o}$ is an order of $\mathfrak{O}$, then $\mathfrak{o}$ is decomposable if and only if $h \neq k$ implies that* $(\mathfrak{O}_h \cap \mathfrak{o}) + (\mathfrak{O}_k \cap \mathfrak{o}) = \mathfrak{o}$, $(h,k = 1,2, \ldots, p)$.

Proof. Suppose that

$$\mathfrak{o} = \sum_{k=1}^{p} \mathfrak{o}_k,$$

where each $\mathfrak{o}_k$ is an order of $\mathfrak{O}_k$. Then evidently,

$$\mathfrak{O}_h \cap \mathfrak{o} = \sum_{\substack{k=1 \\ k \neq h}}^{p} \mathfrak{o}_k,$$

so that $h \neq k$ implies that $(\mathfrak{O}_h \cap \mathfrak{o}) + (\mathfrak{O}_k \cap \mathfrak{o}) = \mathfrak{o}$.

Conversely, if $h \neq k$ implies that $(\mathfrak{O}_h \cap \mathfrak{o}) + (\mathfrak{O}_k \cap \mathfrak{o}) = \mathfrak{o}$, and if we set $\mathfrak{q}_h = \mathfrak{O}_h \cap \mathfrak{o}$, then, in $\mathfrak{o}$,

$$(0) = \bigcap_{h=1}^{p} \mathfrak{q}_h \qquad \text{(direct intersection)}$$

so that

$$\mathfrak{o} = \sum_{k=1}^{p} \mathfrak{o}_k$$

where

$$\mathfrak{o}_k = \bigcap_{\substack{h=1 \\ h \neq k}}^{p} \mathfrak{q}_h \subseteq \mathfrak{O}_k .$$

COROLLARY. *If* $\mathfrak{o}$ *and* $\mathfrak{o}'$ *are orders of* $\mathfrak{D}$, *if* $\mathfrak{o} \subseteq \mathfrak{o}'$ *and if* $\mathfrak{o}$ *is decomposable, then* $\mathfrak{o}'$ *is decomposable.*

Proof. $h \neq k$ implies that

$$1 \in \mathfrak{o} = (\mathfrak{D}_h \cap \mathfrak{o}) + (\mathfrak{D}_k \cap \mathfrak{o}) \subseteq (\mathfrak{D}_h \cap \mathfrak{o}') + (\mathfrak{D}_k \cap \mathfrak{o}')$$

Let us suppose that for $k < q$ $(0 < q < p)$, $\mathfrak{D}_k = \mathfrak{R}_k$, i.e. $\mathfrak{D}_k \cong \mathfrak{D}/\mathfrak{D}_k$ is a field, while for $k > q$, $\mathfrak{D}_k \subset \mathfrak{R}_k$. We consider a commutative ring with unity element in which every regular element is invertible as a Dedekind ring with no relevant prime ideals.

THEOREM 16. *If for each* $k = 1,2, \ldots, p$, $\mathfrak{o}_k$ *is an order of* $\mathfrak{D}_k$, *if for each* $k < q$, $\mathfrak{o}_k$ *is a Dedekind domain and if for* $k > q$, $\mathfrak{o}_k = \mathfrak{D}_k$, *then*

$$\mathfrak{o} = \sum_{k=1}^{p} \mathfrak{o}_k$$

is a Noetherian Dedekind ring, and every order of $\mathfrak{D}$ *that is a Noetherian Dedekind ring may be obtained in this way. If* $\mathfrak{o}_k \subset \mathfrak{D}_k$ *for each* $k \leqslant s$ $(0 < s < q)$ *and* $\mathfrak{o}_k = \mathfrak{D}_k$ *for* $k > s$, *and if, for each* $k \leqslant s$, $\mathfrak{n}_k$ *is the kernel of* Φ_k *in* $\mathfrak{o}$, *then* $\{\mathfrak{n}_k\}_{k<s}$ *is the set of associates of relevant prime ideals of* $\mathfrak{o}$ *and any ideal of* $\mathfrak{o}$ *that contains* $\mathfrak{n}_k$ *properly* $(k < s)$ *must be regular.*

Proof. Let us assume first of all that for each $k < p$, $\mathfrak{o}_k$ is an order of $\mathfrak{D}_k$ obeying the conditions of the theorem. Then,

$$\mathfrak{o} = \sum_{k=1}^{p} \mathfrak{o}_k$$

is an order of $\mathfrak{D}$, and since each $\mathfrak{o}_k$ is fully integrally closed in $\mathfrak{D}_k$, $\mathfrak{o}$ is fully integrally closed in $\mathfrak{D}$. Furthermore, it is evident that $\mathfrak{o}$ is a Noetherian. Let $\mathfrak{a}$ be a regular ideal of $\mathfrak{o}$, $\mathfrak{a} = \mathfrak{a}_1 + \mathfrak{a}_2 + \ldots + \mathfrak{a}_p$, where $\mathfrak{a}_k$ is an ideal of $\mathfrak{o}_k$. Since $\mathfrak{a}$ contains a regular element a of $\mathfrak{o}$ and since $a = a_1 + a_2 + \ldots + a_p$, where a_k is a regular element of $\mathfrak{o}_k$, each $\mathfrak{a}_k$ is a regular ideal of $\mathfrak{o}_k$. Then,

$$\frac{\mathfrak{o}}{\mathfrak{a}} = \sum_{k=1}^{p} \frac{\mathfrak{a} + \mathfrak{o}_k}{\mathfrak{a}} \cong \sum_{k=1}^{p} \frac{\mathfrak{o}_k}{\mathfrak{a} \cap \mathfrak{o}_k} = \sum_{k=1}^{p} \frac{\mathfrak{o}_k}{\mathfrak{a}_k} \qquad \text{(direct sums)}$$

and from the definition of the $\mathfrak{o}_k$, the descending chain condition is valid for the ideals of $\mathfrak{o}_k/\mathfrak{a}_k$, so that it is also valid for the ideals of $\mathfrak{o}/\mathfrak{a}$. Therefore, $\mathfrak{o}$ is a Dedekind ring.

Conversely, let us assume that $\mathfrak{o}$ is an order of $\mathfrak{D}$ and that $\mathfrak{o}$ is a Noetherian Dedekind ring. For each $k < p$, set $\mathfrak{n}_k = \mathfrak{R}_k \cap \mathfrak{o}$ and $\mathfrak{q}_k = \mathfrak{D}_k \cap \mathfrak{o}$. The ideals $\mathfrak{n}_k$ are the only non-regular prime ideals of $\mathfrak{o}$, $h \neq k$ implies that $\mathfrak{n}_h \not\subseteq \mathfrak{n}_k$ and each $\mathfrak{q}_k$ is $\mathfrak{n}_k$-primary.

To prove that $\mathfrak{o}$ is decomposable, by Lemma 9, all we have to show is that $h \neq k$ implies that $\mathfrak{q}_h + \mathfrak{q}_k = \mathfrak{o}$ or equivalently, that $\mathfrak{n}_h + \mathfrak{n}_k = \mathfrak{o}$. If $h \neq k$ and if $\mathfrak{n}$ is a prime ideal of $\mathfrak{o}$ containing $\mathfrak{n}_h + \mathfrak{n}_k$, then evidently, $\mathfrak{n}$ cannot be

contained in any $\mathfrak{n}_j$, and furthermore, $\mathfrak{n}$ cannot be contained in any relevant prime ideal $\mathfrak{p}$ of $\mathfrak{o}$, for then, by Theorem 4, $\mathfrak{n}_h$ and $\mathfrak{n}_k$ would both be contained in the associate $\mathfrak{p}'$ and $\mathfrak{p}$ and would therefore be equal to $\mathfrak{p}'$, a contradiction. Therefore, $\mathfrak{n} = \mathfrak{o}$ and $\mathfrak{n}_h + \mathfrak{n}_k = \mathfrak{o}$.

The associates of relevant prime ideals of $\mathfrak{o}$ must be amongst the $\mathfrak{n}_k$. If $\mathfrak{n}_k$ is the associate of a relevant prime ideal $\mathfrak{p}$ of $\mathfrak{o}$, then $\mathfrak{n}_k$ is the only $\mathfrak{n}_k$-primary ideal of $\mathfrak{o}$ (see remarks following Theorem 7) so that

$$\mathfrak{o}_k = \Phi_k(\mathfrak{o}) \cong \mathfrak{o}/\mathfrak{n}_k$$

is an integral domain and not a field since it contains the non-trivial ideal $\Phi_k(\mathfrak{p})$ and therefore, $k \leqslant s$. Furthermore, if $\mathfrak{a}$ is an ideal of $\mathfrak{o}$ that contains $\mathfrak{n}_k$ properly, then $\mathfrak{a}$ is regular, for if $\mathfrak{a}$ were not regular, then $\mathfrak{O}\mathfrak{a}$ would be a proper ideal of $\mathfrak{O}$ and would therefore be contained in some $\mathfrak{N}_j$ and then

$$\mathfrak{n}_k \subset \mathfrak{a} \subseteq \mathfrak{O}\mathfrak{a} \cap \mathfrak{o} \subseteq \mathfrak{n}_j,$$

a contradiction. Then, $\mathfrak{a}$ may be expressed as a product of relevant prime ideals of $\mathfrak{o}$ having $\mathfrak{n}_k$ as associate (standard decomposition of $\mathfrak{a}$) so that $\Phi_k(\mathfrak{a})$ may be expressed as a product of prime ideals of $\mathfrak{o}_k$ and therefore, $\mathfrak{o}_k$ is a Dedekind domain.

If $\mathfrak{n}_k$ is not the associate of a relevant prime ideal of $\mathfrak{o}$, then $\mathfrak{n}_k$ is a maximal proper ideal of $\mathfrak{o}$ so that $\mathfrak{o}_k \cong \mathfrak{o}/\mathfrak{q}_k$ is a completely primary ring and therefore, $\mathfrak{o}_k = \mathfrak{O}_k$ so that $k > s$.

In the following corollaries, $\mathfrak{o}$ is a Noetherian Dedekind ring having $\mathfrak{O}$ as full ring of quotients and we adopt the notation developed in Theorem 16.

COROLLARY 1. *If for $k \leqslant s$,*

$$\{\mathfrak{p}_i\}_{i \in I_k}$$

is the set of relevant prime ideals of $\mathfrak{o}$ that have $\mathfrak{n}_k$ as associate, then

$$\{\Phi_k(\mathfrak{p}_i)\}_{i \in I_k}$$

is the set of relevant prime ideals of $\mathfrak{o}_k$ and for each $i \in I_k$, $\mathfrak{p}_i = \Phi_k^{-1}\Phi_k(\mathfrak{p}_i)$.

COROLLARY 2. *If for each $i \in I_k$, V_i is the valuation of $\mathfrak{O}$ determined by $\mathfrak{p}_i$ and V_i' is the valuation of $\mathfrak{O}_k$ determined by $\Phi_k(\mathfrak{p}_i)$, then V_i' is the projection of V_i by Φ_k.*

COROLLARY 3. *If $\mathfrak{o}'$ is an order of $\mathfrak{O}$ and if $\mathfrak{o} \subseteq \mathfrak{o}'$, then $\mathfrak{o}'$ is also a Noetherian Dedekind ring.*

Proof. By the Corollary of Lemma 9, $\mathfrak{o}'$ is also decomposable, and since for each k, $\Phi_k(\mathfrak{o}) \subseteq \Phi_k(\mathfrak{o}')$, for each $k > s$, $\Phi_k(\mathfrak{o}') = \mathfrak{O}_k$ and by a theorem of MacLane and Schilling (**6**, Lemma 37), for $k \leqslant s$, $\Phi_k(\mathfrak{o}')$ is a Dedekind domain, so that by Theorem 16, $\mathfrak{o}'$ is a Noetherian Dedekind ring. It is evident that this Corollary is a generalization of the above mentioned Theorem of MacLane and Schilling.

References

1. G. Birkhoff, *Lattice Theory*, Amer. Math. Soc. Colloquium Publications, 25.
2. L. Fuchs, *The generalization of the valuation theory*, Duke Math. J., *18* (1951), 19–26.
3. W. Krull, *Über die Zerlegung der Hauptideale in all-gemeinen Ringen*, Math. Ann., *105* (1931), 1–14.
4. ——, *Idealtheorie* (1935; Chelsea, 1948).
5. D. G. Northcott, *Ideal Theory* (Cambridge, 1953).
6. O. F. G. Schilling, *The Theory of Valuations*, Math. Surveys of the Amer. Math. Soc., IV (1950).
7. B. L. van der Waerden, *Moderne Algebra* (1937; 2nd ed. New York, 1943).

Université de Montréal

ANNOUNCEMENT

The Canadian Mathematical Congress announces the retirement of Professor H. S. M. Coxeter from the position of Editor-in-Chief of the Canadian Journal of Mathematics. His place will be taken by Professor G. F. D. Duff and contributors are therefore requested to address correspondence to the new Editor,

c/o Department of Mathematics,
University of Toronto,
Toronto 5, Ontario,
Canada.